



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17		(11) International Publication Number: WO 98/39446
A2		(43) International Publication Date: 11 September 1998 (11.09.98)
(21) International Application Number: PCT/US98/04482 (22) International Filing Date: 6 March 1998 (06.03.98)		
(30) Priority Data: 60/040,162 7 March 1997 (07.03.97) US 60/040,333 7 March 1997 (07.03.97) US 60/038,621 7 March 1997 (07.03.97) US 60/040,161 7 March 1997 (07.03.97) US 60/040,626 7 March 1997 (07.03.97) US 60/040,334 7 March 1997 (07.03.97) US 60/040,336 7 March 1997 (07.03.97) US 60/040,163 7 March 1997 (07.03.97) US 60/043,580 11 April 1997 (11.04.97) US 60/043,568 11 April 1997 (11.04.97) US (Continued on the following page)		
(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).		
(72) Inventors; and (75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BEDNARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda, MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hia [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).		
(74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US).		
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).		
Published Without international search report and to be republished upon receipt of that report. With an indication in relation to a deposited microorganism furnished under Rule 13bis separately from the description. Date of receipt by the International Bureau: 06 April 1998 (06.04.1998)		

(54) Title: 70 HUMAN SECRETED PROTEINS

(57) Abstract

The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.

60/043,314	11 April 1997 (11.04.97)	US	60/047,598	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US
60/043,569	11 April 1997 (11.04.97)	US	60/047,613	23 May 1997 (23.05.97)	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,311	11 April 1997 (11.04.97)	US	60/047,582	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,671	11 April 1997 (11.04.97)	US	60/047,596	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,674	11 April 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,669	11 April 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,312	11 April 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05.97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,313	11 April 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05.97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,672	11 April 1997 (11.04.97)	US	60/047,599	23 May 1997 (23.05.97)	US	60/056,636	22 August 1997 (22.08.97)	US
60/043,315	11 April 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05.97)	US	60/056,874	22 August 1997 (22.08.97)	US
60/043,578	11 April 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05.97)	US	60/056,910	22 August 1997 (22.08.97)	US
60/043,576	11 April 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08.97)	US
60/043,670	11 April 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05.97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,594	23 May 1997 (23.05.97)	US	60/056,845	22 August 1997 (22.08.97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,589	23 May 1997 (23.05.97)	US	60/056,892	22 August 1997 (22.08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05.97)	US	60/056,632	22 August 1997 (22.08.97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,876	22 August 1997 (22.08.97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,881	22 August 1997 (22.08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,909	22 August 1997 (22.08.97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,875	22 August 1997 (22.08.97)	US
60/047,503	23 May 1997 (23.05.97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,862	22 August 1997 (22.08.97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22.08.97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22.08.97)	US
60/047,584	23 May 1997 (23.05.97)	US	60/056,630	22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
60/047,500	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
60/047,587	23 May 1997 (23.05.97)	US	60/056,662	22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047,492	23 May 1997 (23.05.97)	US	60/056,872	22 August 1997 (22.08.97)	US			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

70 Human Secreted Proteins

Field of the Invention

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and
5 their production.

Background of the Invention

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or
10 organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum
15 (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

20 Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or
25 secretory granules) until exocytosis is triggered. Similarly, proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins include
30 the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent medical disorders by using
35 secreted proteins or the genes that encode them.

Summary of the Invention

The present invention relates to novel polynucleotides and the encoded polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting disorders related to the polypeptides, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying binding partners of the polypeptides.

Detailed Description

Definitions

The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:X was often generated by overlapping sequences contained in multiple clones (contig

analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 12301 Park Lawn Drive, Rockville, Maryland 20852, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM sodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH₂PO₄; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 µg/ml salmon sperm blocking DNA; followed by washes at 50°C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

5 The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be
10 single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability
15 or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

 The polypeptide of the present invention can be composed of amino acids joined
20 to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs,
25 as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be
30 branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a
35 nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine,

formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS -
5 STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990);
10 Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting
15 activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present
20 invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

25 Polynucleotides and Polypeptides of the Invention

FEATURES OF PROTEIN ENCODED BY GENE NO: 1

The translation product of Gene NO: 1 shares sequence homology with alpha-L-fucosidase which is thought to be important as a lysosomal enzyme that hydrolyzes
30 fucose from fucoglycoconjugates. (See Accession No. gi/178409.) Lysosome fructosidase is involved in certain lysosome storage diseases. (See Biochem. Biophys. Res. Commun., 164(1):439-445 (1989).) Fucosidosis, an autosomal recessive lysosomal storage disorder characterized by progressive neurological deterioration and mental retardation. The disease results from deficient activity of alpha-L-fucosidase, a
35 lysosomal enzyme that hydrolyzes fucose from fucoglycoconjugates. This gene likely encodes a novel fucosidase isoenzyme. Based on homology, it is likely that the translated product of this gene is also involved in lysosome catabolism of molecules and

that aberrations in the concentration and/or composition of this product may be causative in lysosome storage disorders. Preferred polypeptide fragments comprise the amino acid sequence PGHLLPHKWENC (SEQ ID NO: 257).

Gene NO: 1 is expressed primarily in stromal cells, and to a lesser extent in human fetal kidney and human tonsils.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, fucosidosis and other lysosome storage disorders. Similarly, polypeptides and antibodies directed to the polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues of cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., stromal cells, kidney, tonsils, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 1 to alpha-L-fucosidase indicates that polypeptides and polynucleotides corresponding to Gene NO: 1 are useful for the treatment of fucosidosis and general lysosomal disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 134 as residues: Met-1 to Leu-6, Thr-32 to Glu-39, Lys-80 to Lys-85, and Met-90 to Pro-96.

FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The translation product of Gene No. 2 shares sequence homology with stromal cell-derived factor-2 (SDF-2) which is a novel secreted factor. See, for example, Gene. 176(1-2):211-214, (1996, Oct. 17.) The amino acid sequence of SDF-2 shows similarity to yeast dolichyl phosphate-D-mannose:protein mannosyltransferases, Pmt1p [Strahl-Bolsinger et al. Proc. Natl. Acad. Sci. USA 90, 8164-8168 (1993)] and Pmt2p [Lussier et al. J. Biol. Chem. 270, 2770-2775 (1995)], whose activities have not been detected in higher eukaryotes. Based on the sequence similarity, the translation product of this gene is expected to share certain biological activities with SDF-2, Pmt1p and Pmt2p.

Gene NO: 2 is expressed primarily in immune system tissue and cancerous tissues, such as liver hepatoma, human B-cell lymphoma, spleen in a patient suffering

from chronic lymphocytic leukemia, hemangiopericytoma, pharynx carcinoma, breast cancer, thyroid, bone marrow, osteoblasts and to a lesser extent in a few other tissues such as kidney pyramids.

Therefore, polynucleotides or polypeptides of the invention are useful as
5 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of the diseases and conditions which include, but are not limited to, disorders in kidney, liver, and immune organs, particularly cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
10 type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, liver, thyroid, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, spleen, B-cells, pharynx, thyroid, mammary tissue, bone marrow, osteoblasts and kidneys, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma,
15 urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 2 to stromal cell-derived
20 factor-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 2 are useful for diagnosis and therapeutic treatment of disorders in kidney, liver, and immune organs since stromal cells play important role in organ function. Stroma carries the blood supply and provides support for the growth of parenchymal cells and is therefore crucial to the growth of a neoplasm. Nucleic acids of the present invention
25 comprise, but preferably do not consist of, and more preferably do not comprise, SEQ ID NO: 3 from US Patent No. 5,576,423, incorporated herein by reference, and shown herein as SEQ ID NO: 258).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
135 as residues: His-56 to Gly-65, Ala-74 to Ser-80, Ile-84 to Pro-97, Leu-124 to Glu-
30 129, Glu-135 to Asp-143, Gly-175 to Ser-180, and Ala-194 to Thr-199.

FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The translation product of Gene NO: 3 shares sequence homology with LZIP-1, LZIP-2 and other leucine zipper proteins, which are thought to be important in nucleic
35 acid binding. This gene has been reported in Mol. Cell. Biol. 17 (9), 5117-5126 (1997) as "Luman". Luman is a cyclic AMP response element (CRE)-binding protein/activating transcription factor 1 protein of the basic leucine zipper superfamily. It binds CREs in

vitro and activates CRE-containing promoters when transfected into COS7 cells. The complete amino acid sequence of Luman reported in Mol. Cell. Biol. 17 (9): 5117-5126 (1997) is:

MELELDAGDQDLLAFLLEESGDLGTAPDEAVRAPLDWALPLSEVPSDWEVDDL
 5 CSLLSPPASLNILSSSNPCLVHHHTYSLPRETVSMDLESESCRKEGTQMT PQH
 MEELAEQEIARLVLTDEEKSLLLEKEGLILPETLPLTKTEEQILKRVRRKIRNKRSA
 QESRRKKKKVYVGGLESRLKYTAQNMELQNKVQLLEEQLNSLLDQLRKLQAM
 VIEISNKTSSSSTCILVLLVSFCLLLVPAMYSSDTRGSLPAEHGVLSRQLRALPSE
 DPYQLELPALQSEVPKDSHTQWLDGSDCVLQAPGNTSCLLHYMPQAPSAEPPL
 10 EWPFPDLSS EPLCRGPILPLQANLTRKGGWLPTGSPSVILQDRYSG (SEQ ID
 N:259).

Gene NO: 3 is expressed primarily in apoptotic T-cells and Soares senescent cells and to a lesser extent in multiple tissues and cell types, including, multiple sclerosis tissue, and hippocampus.

15 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. Similarly, polypeptides and antibodies directed to these
 20 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and transplantation, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., multiple sclerosis tissue, hippocampus, bone marrow and cancerous and wounded
 25 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 3 to leucine zipper nucleic
 30 acid binding proteins indicates that polypeptides and polynucleotides corresponding to Gene NO: 3 are useful for diagnosis and treatment of immunologically mediated disorders, transplantation, immunodeficiency, and tumor necrosis. The secreted nucleic acid binding protein in the apoptotic tissues may be involved in the disposal of the DNA released by apoptotic cells. Furthermore, the studies conducted in support of Luman
 35 suggest that the translation product of this gene may be used to identify transcriptional regulation elements which in turn are useful in modulation of immune function.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 136 as residues: Asn-7 to Ser-12, Tyr-32 to Gly-38, Pro-55 to Tyr-60, Glu-70 to Thr-76, and Pro-104 to Leu-110.

5 FEATURES OF PROTEIN ENCODED BY GENE NO: 4

The translation product of Gene NO: 4 shares sequence homology with a number of tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene, CO-029 tumor associated antigen protein, CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, which are thought to be important in development
10 of cancer, immune system development and functions.

Gene NO: 4 is expressed primarily in cancers of several different tissues and to a lesser extent in normal tissue like prostate, skin and kidney.

Therefore, polynucleotides or polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers and disorders of the immune system, prostate and kidney. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell
20 type(s). For a number of disorders of the above tissues or cells, particularly of the kidney, skin, prostate and immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., kidney, skin and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
25 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 4 to tetraspan transmembrane surface molecules such as human metastasis tumor suppressor gene.
30 CO-029 tumor associated antigen protein, CD53 hematopoietic antigen, human membrane antigen TM4 superfamily protein, metastasis controlling peptide, and human CD9 sequence, indicates that polypeptides and polynucleotides corresponding to Gene NO: 4 are involved with the cellular control of growth and differentiation. Therefore, the translation product of this gene is believed to be useful for diagnosis and treatment
35 of neoplasia and disorders of the kidney, skin and prostate. For example, recombinant protein can be produced in transformed host cells for diagnostic and prognostic applications. Alterations in the protein sequence are indicative of the presence of

malignant cancer, or of a predisposition to malignancy, in a subject. Gene therapy can be used to restore the wild-type gene product to a subject. Additionally, the antibodies are a useful tool for the identification of hematopoietic neoplasms, and may prove helpful for identifying morphologically poorly defined cells. Moreover, this protein can be used to isolate cognate receptors and ligands and identify potential agonists and antagonists using techniques known in the art. The protein also has immunomodulatory activity, regulates hematopoiesis and stimulates growth and regeneration as a male/female contraceptive, increases fertility depending on activin and inhibin like activities. Other uses are as a chemotactic agent for lymphocytes, treatment of coagulation disorders, an anti-inflammatory agent, an antimicrobial or analgesic and as a modulator of behavior and metabolism. The DNA can be used in genetic diagnosis or gene therapy, and for the production of recombinant protein. It can also be used to identify protein expressing cells, isolate related sequences, prepare primers for genetic fingerprinting and generate anti-protein or anti-DNA antibodies. In addition, residues 1-71, in the translation product for this gene are believed to be the extracellular domain. Thus, polypeptide comprising residues 1-71 or derivatives (including fragments) or analogs thereof, are useful as a soluble polypeptide which may be routinely used therapeutically to antagonize the activities of the receptor.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 137 as residues: Lys-118 to Phe-127, Asn-145 to Ala-160, and Thr-177 to Val-188.

FEATURES OF PROTEIN ENCODED BY GENE NO: 5

Gene NO: 5 is expressed primarily in human testes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the testes including cancer and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of Gene NO: 5 indicates that the protein product of this gene is useful for treatment/diagnosis of diseases of the testes, particularly testicular cancer since expression is observed primarily in the testes.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 5 138 as residue: Gly-22 to Gln-30.

FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The translation product of Gene NO: 6 shares sequence homology with GALNS (N-acetylgalactosamine 6-sulphatase) which is thought to be important in the storage of
10 the glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. See Genbank accession no. gil618426. Based on the sequence similarity, the translation product of this gene is expected to share biological activities with GALNS.

Gene NO: 6 is expressed primarily in human bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as
15 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate, e.g., Morquio A syndrome. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
20 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly involving cell storage disorder, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
25 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology of Gene NO: 6 to N-acetylgalactosamine 6-sulphatase indicates that polypeptides and polynucleotides corresponding to Gene
30 NO: 6 are useful for the treatment and diagnosis of storage disorders of glycosaminoglycans, keratan sulfate and chondroitin 6-sulfate. Such disorders are known in the art and include, e.g., Morquio A syndrome which is caused by an error of mucopolysaccharide metabolism with excretion of keratan sulfate in urine. Morquio A syndrome is characterized by severe skeletal defects with short stature, severe deformity
35 of spine and thorax, long bones with irregular epiphyses but with shafts of normal length, enlarged joints, flaccid ligaments, and waddling gait; autosomal recessive inheritance.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 139 as residues: Gly-29 to Pro-36 and Glu-57 to Leu-64.

FEATURES OF PROTEIN ENCODED BY GENE NO: 7

5 The translation product of Gene NO: 7 shares sequence homology with carboxy peptidase E and H (carboxypeptidase E is thought to be important in the biosynthesis of numerous peptide hormones and neurotransmitters). The translation product of this gene also shares sequence homology with bone-related carboxypeptidase "OSF-5" from the mouse. See European patent application EP-588118-A. Based on the sequence
10 similarity to OSF-5, the translation product of this gene will hereinafter sometimes be referred to as "human-OSF-5" or "hOSF-5".

Gene NO: 7 is expressed primarily in tumor cell lines derived from connective tissues including chondrosarcoma, synovial sarcoma, Wilm's tumor and rhabdomyosarcoma and to a lesser extent in a myeloid progenitor cell line, bone
15 marrow, and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, various cancers involving the skeletal system and connective tissues in
20 general, in particular at cartilage interfaces. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal system and various other tumor tissues, expression of this gene at significantly higher or lower levels may routinely be
25 detected in certain tissues (e.g., connective tissues and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

30 The restricted tissue distribution and homology of Gene NO: 7 to carboxypeptidase E and mouse OSF-5 indicates that polypeptides and polynucleotides corresponding to Gene NO: 7 are for processing of peptides to their mature form that may have various activities similar to the activities of neuropeptides but in the periphery. In addition the abundance of expression in cancer tissues indicates that
35 aberrant expression and subsequent processing may play a role in the progression of malignancies, e.g., growth factor and/or adhesion factor activities. In particular, the expression of this gene is restricted to connective tissues and embryonic tissues.

Furthermore, it is overexpressed in cancers of these same tissues (i.e., in sarcomas). Moreover, hOSF-5 shares very strong sequence similarity with mOSF-5 which is a known bone growth factor and is thought to be useful in obtaining products for the diagnosis and treatment of bone metabolic diseases, e.g., osteoporosis and Paget's disease. Like OSF-5, the translation product of this gene is believed to be a bone-specific carboxypeptidase which acts as an adhesion molecule/growth factor and takes part in osteogenesis at the site of bone induction. hOSF-5 can, therefore, be used to treat bone metabolic diseases, osteoporosis, Paget's disease, osteomalacia, hyperostosis or osteopetrosis. Furthermore, hOSF-5 can be used to stimulate the regeneration of bone at the site of mechanical damage, e.g., accidentally or surgically caused fractures.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 140 as residues: Leu-24 to Val-30, Ala-89 to Lys-94, Phe-150 to Trp-157, Leu-162 to Asp-167, Asp-187 to Ser-199, His-241 to Asp-254, and Pro-362 to Asp-376.

FEATURES OF PROTEIN ENCODED BY GENE NO: 8

Gene NO: 8 is expressed primarily in bone marrow, and to a lesser extent in an erythroleukemia cell line.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hematological disorders including cancer and anemia. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematologic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 8 are useful as a growth factor for hematopoietic stem cells or progenitor cells, e.g., in the treatment of bone marrow stem cell loss in chemotherapy patients and in the treatment of kidney disease.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 141 as residues: Gly-30 to Lys-35.

FEATURES OF PROTEIN ENCODED BY GENE NO: 9

Gene NO: 9 is expressed primarily in neutrophils.

Therefore, polynucleotides or polypeptides of the invention are useful as
5 reagents for differential identification of the cell type present in a biological sample and
for diagnosis of diseases and conditions which include, but are not limited to,
inflammatory diseases. Similarly, polypeptides and antibodies directed to these
polypeptides are useful in providing immunological probes for differential identification
of the cell type indicated. For a number of disorders of the above tissues or cells,
10 particularly of the immune system, expression of this gene at significantly higher or
lower levels may routinely be detected in certain tissues or cell types (e.g., neutrophils,
bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum,
plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
an individual having such a disorder, relative to the standard gene expression level, i.e.,
15 the expression level in healthy tissue or bodily fluid from an individual not having the
disorder.

The tissue distribution indicates that polypeptides and polynucleotides
corresponding to Gene NO: 9 are useful for immune modulation or as a growth factor
to stimulate neutrophil differentiation or proliferation that may be useful in the treatment
20 of neutropenia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
142 as residues: Thr-22 to Pro-37.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

25 Gene NO: 10 is expressed primarily in the epidermis.

Therefore, polynucleotides or polypeptides of the invention are useful as
reagents for differential identification of the tissue(s) or cell type(s) present in a
biological sample and for diagnosis of diseases and conditions which include, but are
not limited to, diseases of the epidermis such as psoriasis or eczema or may be involved
30 in the normal proliferation or differentiation of the epithelial cells or fibroblasts
constituting the skin. Similarly, polypeptides and antibodies directed to these
polypeptides are useful in providing immunological probes for differential identification
of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
particularly of the skin, expression of this gene at significantly higher or lower levels
35 may routinely be detected in certain tissues (e.g., epidermis and cancerous and
wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal
fluid) or another tissue or cell sample taken from an individual having such a disorder.

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 10 are useful for diagnosis and treatment of skin conditions and as an aid in the healing of various epidermal injuries including wounds, and
5 diabetic ulcers.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 143 as residues: Ser-3 to Ser-9 and Trp-27 to Glu-32.

10 FEATURES OF PROTEIN ENCODED BY GENE NO: 11

The translation product of Gene NO: 11 shares sequence homology with phosphatidylcholine 2-acylhydrolase (PLA2). See, for example, Genbank accession no. gi190004. PLA2 is involved in inflammation, where it is responsible for the conversion of cell membrane phospholipids into arachidonic acid. Arachidonic acid in
15 turn feeds into both the lipoxygenase and cyclooxygenase pathways to produce leukotrienes (involved in chemotaxis, vasoconstriction, bronchoconstriction, and increased vascular permeability) and prostaglandins (responsible for vasodilation, potentiate edema, and increased pain). Diseases in which PLA2 is implicated as a major factor include rheumatoid arthritis, sepsis, ischemia, and thrombosis. The inventors
20 refer to the translation product of this gene as PLA2-like protein based on the sequence similarity. Furthermore, owing to the sequence similarity PLA2 and PLA2-like protein are expected to share certain biological activities.

Gene NO: 11 is expressed primarily in human cerebellum and in T-cells.

Therefore, polynucleotides or polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s)
30 or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cerebellum and Purkinje cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, bone marrow, T-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
35 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 11 are useful for diagnosis and treatment of cerebellum disorders, rheumatoid arthritis, sepsis, ischemia, and thrombosis. This gene is also useful as a chromosome marker. It is believed to map to Chr.15, D15S118-D15S123.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 12

Gene NO: 12 is expressed primarily in highly vascularized tissues such as placenta, uterus, tumors, fetal liver, fetal spleen and also in the C7MCF7 cell line treated with estrogen.

10

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, endometriosis, endometritis, endometrial carcinoma, primary hepatocellular carcinoma, and spleen-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the endometrium, liver and spleen, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endometrium, liver, and spleen, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15

20

25

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 12 are useful for diagnosis and treatment of diseases of the endometrium (such as endometrial carcinoma, endometriosis, and endometritis), liver diseases (such as primary hepatocellular carcinoma), and spleen-related diseases.

SEQ ID NO: 145 as residues: Ala-29 to Leu-35, Leu-50 to Ser-57, Glu-96 to Glu-105, Asp-140 to Asp-148, and Asn-191 to Ser-197.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 13

Gene NO: 13 is expressed primarily in B cell lymphoma and to a lesser extent in other tissues.

35

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B cell lymphoma; hematopoietic disorders; immune dysfunction.

Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may
5 routinely be detected in certain tissues and cell types (e.g., bone marrow and B-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the
10 expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Enhanced expression of this gene product in B cell lymphoma indicates that it may play a role in the proliferation of hematopoietic cells. It is also believed to be involved in the survival and/or differentiation of various hematopoietic lineages. Expression in lymphoma also indicates that it may be involved in other cancers and
15 abnormal cellular proliferation. The tissue distribution, therefore, indicates that polypeptides and polynucleotides corresponding to Gene NO: 13 are useful for the diagnosis and/or therapeutic treatment of hematopoietic disorders, particularly B cell lymphoma. Furthermore, since overexpression of this gene is associated with the development of B cell lymphoma, antagonists of this protein are useful to interfere with
20 the progression of the disease. This protein is useful in assays for identifying such antagonists. Assays for identifying antagonists are known in the art and are described briefly elsewhere herein. Preferred antagonists include antibodies and antisense nucleic acid molecules. Preferred are antagonists which inhibit B-cell proliferation.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 14**

The translation product of Gene NO: 14 shares sequence homology with very low density lipoprotein receptor which is thought to be important in transport of lipoproteins. Owing to the sequence similarity the translation product of this gene is believed to share certain biological activities with VLDL receptors. Assaying such
30 activity may be achieved by assays known in the art and set forth elsewhere herein.

This gene is expressed primarily in human synovium, umbilical vein endothelial cells, CD34+ cells, Jurkat cells, and HL60 cells, and to a lesser extent in thymus, meningioma, hypothalamus, adult testis, and fetal liver and spleen.

Therefore, polynucleotides or polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, atherosclerosis, ataxia malabsorption, vascular damage, hyperlipidemia,

and other cardiovascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular and hematological systems.

- 5 expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., endothelium, thymus meningioma, hypothalamus, testes, liver, and spleen and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,
10 the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- The tissue distribution in the vascular endothelial cells and homology to VLDL receptors indicates that polypeptides and polynucleotides corresponding to Gene NO: 14 are useful for diagnosis and treatment of atherosclerosis, ataxia malabsorption, and
15 hyperlipidemia. These and other factors often result in other cardiovascular diseases. Additionally, the presence of the gene product in cells of blood lineages indicates that it may be useful in hematopoietic regulation and hemostasis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 147 as residues: Pro-39 to Ser-52, Trp-71 to Thr-76, and Pro-94 to His-100.

20

FEATURES OF PROTEIN ENCODED BY GENE NO: 15

- The translation product of Gene NO: 15 shares sequence homology with kallikrein which is thought to be important in blood pressure and renal secretion. Furthermore, this gene has now been characterized as a novel hepatitis B virus X
25 binding protein that inhibits viral replication. See, for example, J. Virol. 72 (3), 1737-1743 (1998).

This gene is expressed primarily in kidney, placenta, lung, aorta and other endothelial cells, caudate nucleus and to a lesser extent in melanocytes, liver, adipose tissue.

- 30 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renovascular hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy. Similarly, polypeptides and antibodies directed to these
35 polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the renovascular or respiratory vascular systems, expression of this gene

at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., kidney, placenta, lung, endothelial cells, melanocytes, liver, and adipose tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from
5 an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to kallikrein indicates that polypeptides and polynucleotides corresponding to Gene NO: 15 are useful for treating renovascular
10 hypertension, renal secretion, electrolyte metabolism, toxemia of pregnancy and hydronephrosis. The protein expression in the organs like kidney, lung and vascular endothelial cells indicates the gene involvement in hemodynamic regulatory functions. The translation product of this gene is also useful in the treatment of viral infection, particularly liver infection, and particularly hepatitis B virus(es).

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 148 as residues: Leu-9 to Asn-15 and Thr-56 to Asp-61.

FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The translation product of Gene NO: 16 shares sequence homology with
20 secretory component protein, immunoglobulins and their receptors which are thought to be important in immunological functions. The amino acid sequence of secretory component protein can be accessed as accession no. pirlA02112, incorporated herein by reference.

Gene NO: 16 is expressed primarily in macrophages, monocytes and dendritic
25 cells and to a lesser extent in placenta and brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation and tumors. Similarly, polypeptides and antibodies directed
30 to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cells (e.g., macrophages, monocytes, dendritic cells, placenta and brain, and cancerous
35 and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a

disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to immunoglobulins and secretory component protein indicates that polypeptides and polynucleotides corresponding to
5 Gene NO: 16 are useful for diagnosis and treatment of inflammation and bacterial infection, and other diseases where immunomodulation would be beneficial.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 149 as residues: Pro-37 to Cys-51, Gln-53 to Cys-60, Asn-99 to Gly-106, Gly-145 to Glu-151, and Ile-159 to Ser-164.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 17

The translation product of Gene NO: 17 is evolutionarily conserved and shares sequence homology with proteins from yeast and *C. elegans*. See, for example, Genbank accession no.gil746540. As is known in the art, strong sequence similarity to
15 a secreted protein from *C. elegans* is predictive of cellular location of human proteins.

Gene NO: 17 is expressed primarily in colon carcinoma cell lines, messangial cells, many tumors like T cell lymphoma, osteoclastoma, Wilm's tumor, adrenal gland tumor, testes tumor, synovial sarcoma, and to a lesser extent in placenta, lung and brain.

20

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rapidly growing/dividing cells such as cancerous tissue, including, colon carcinoma, lymphomas, and sarcomas. Similarly, polypeptides and antibodies directed
25 to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the gastrointestinal, hematological and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, lung, brain, colon, messangial cells, adrenal gland, T-cells, testes, and lymph tissue, and cancerous and wounded tissues) or
30 bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35

The tissue distribution in colon cancer and many other tumors indicates that the polynucleotides and polypeptides of Gene NO: 17 are useful for cancer diagnosis and therapeutic targeting. The extracellular nature may contribute to solid tumor

immunosuppression, angiogenesis and cell growth stimulation. The tissue distribution of this gene in cells of the immune system indicates that polypeptides and polynucleotides corresponding to Gene NO: 17 are useful for treatment, prophylaxis and diagnosis of immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. Its expression predominantly in hematopoietic cells also indicates that the gene could be important for the treatment and/or detection of hematopoietic disorders such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The protein can also be used to enhance or protect proliferation, differentiation and functional activation of hematopoietic progenitor cells such as bone marrow cells, which could be useful for cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The protein may also be useful to increase the proliferation of peripheral blood leukocytes, which could be useful in the combat of a range of hematopoietic disorders including immunodeficiency diseases, leukemia, and septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 150 as residues: Val-131 to Asn-136.

FEATURES OF PROTEIN ENCODED BY GENE NO: 18

The translation product of Gene NO: 18 shares sequence homology with immunoglobulin, which is thought to be important in immunoreactions.

Gene NO: 18 is expressed primarily in macrophage.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, inflammation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., macrophage and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution in macrophages and the weak homology to immunoglobulin indicates that polypeptides and polynucleotides corresponding to Gene

NO: 18 are useful for diagnosing and treating immune response disorders, including inflammation, antigen presentation and immunosurveillance.

FEATURES OF PROTEIN ENCODED BY GENE NO: 19

5 The translation product of Gene NO: 19 shares sequence homology with proline rich proteins which are thought to be important in protein-protein interaction.

 This gene has a wide range of tissue distribution, but is expressed primarily in normal prostate, synovial fibroblasts, brain amygdala depression, fetal bone and fetal cochlea, and to a lesser extent in adult retina, umbilical vein endothelial cells, atrophic
10 endometrium, osteoclastoma, melanocytes, pancreatic carcinoma and smooth muscle.

 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer metastasis, wound healing, tissue repair. Similarly, polypeptides
15 and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skeletal, connective tissues, reproductive and central nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain,
20 prostate, fibroblasts, bone, cochlea, retina, endothelial cells, endometrium, pancreas and smooth muscle, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the
25 disorder.

 The tissue distribution and homology to proline-rich proteins indicates that the protein is a extracellular matrix protein or an ingredient of bodily fluid. Polypeptides and polynucleotides corresponding to Gene NO: 19 are useful for cancer metastasis intervention, tissue culture additive, bone modeling, wound healing and tissue repair.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 20

 Gene NO: 20 is expressed primarily in prostate cancer, leukocytes, meningioma, adult liver, pancreas, brain, and to a lesser extent in lung.

 Therefore, polynucleotides or polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate cancers. Similarly, polypeptides and antibodies directed to these

polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate and brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., prostate, leukocytes, meningioma, liver, brain, pancreas and lung, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5
10 Prostate cancer cell lines are known to be responsive to estrogen and androgen. The protein expression of Gene NO: 20 appears to be influenced by both estrogen and androgen levels. The prostate cancer tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 20 are is useful in the intervention and detection of prostate hyperplasia and prostate cancer.

15

FEATURES OF PROTEIN ENCODED BY GENE NO: 21

- The translation product of Gene NO: 21 is identical to the human wnt-7a gene. Wnt-7a is a secreted signaling molecule, thought to be important in signaling and the regulation of cell fate and pattern formation during embryogenesis. Specifically, knock out studies in mice have demonstrated that wnt7a plays a critical role in the development of the dorsal-ventral patterning in the developing limb, and to a lesser extent plays a role in the development of anterior-posterior patterning. Overexpression of wnt7a can induce transformation of cultured mammary cells, suggesting that it is an oncogene.

- 20
25 Expression of Gene NO: 21 has only been observed in testes. Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, testicular cancer; abnormal limb development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the testes or developing embryo. For a number of disorders of the above tissues or cells, particularly of the developing embryo, expression of this gene at significantly higher or lower levels may routinely be detected in the developing embryo or amniotic fluid taken from a pregnant individual and compared relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Also, expression of this gene at significantly higher or lower levels may routinely be detected in the testes of patient suffering from testicular cancer and compared relative to the
- 30
35

standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to mouse *wnt7a* indicates that polypeptides and polynucleotides corresponding to Gene NO: 21 are useful to restore abnormal limb development in an affected individual. Furthermore, its oncogenic potential and tissue distribution indicates that it could serve as a diagnostic for testicular cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 154 as residues: Gly-22 to Arg-28.

10

FEATURES OF PROTEIN ENCODED BY GENE NO: 22

Gene NO: 22 is expressed primarily in fetal liver/spleen, breast, testes and placenta and to a lesser extent in brain, and a series of cancer tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, brain diseases, male infertility, and disposition to pregnant miscarriages. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoietic system, and sexual organs, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, testes, placenta, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution of this gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 22 are useful as a marker for non-differentiated, dividing cells and hence could serve as an oncogenic marker. Its high expression in fetal liver, suggests an involvement in hematopoiesis and/or the immune system. Hence it is useful as a factor to enhance an individual's immune system, e.g., in individuals with immune disorders. It is also thought to affect the survival, proliferation, and differentiation of a number of hematopoietic cell lineages, including hematopoietic stem cells. Its disruption, e.g., mutation or altered expression, may also be a marker of immune disorder. Its expression in the testes, suggests it may be important in controlling male fertility. Expression of this gene in breast further reflects a

role in immune function and immune surveillance (breast lymph node). This gene is believed to be useful as a marker for breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 155 as residues: Gln-57 to Lys-70 and Ala-91 to Pro-100.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 23

Gene NO: 23 is expressed primarily in bone marrow and brain (whole and fetal).

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological, immune and hematopoietic disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the central nervous and hematopoietic systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., bone marrow, brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 23 are useful in the diagnosis and treatment of disorders related to the central nervous system (e.g. neuro-degenerative conditions, trauma, and behavior abnormalities) and hematopoiesis. In addition, the expression in fetal brain indicates a role for this gene product in diagnosis of predisposition to developmental defects of the brain.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 156 as residues: Thr-23 to Tyr-29.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 24

Gene NO: 24 is expressed primarily in smooth muscle, placenta, prostate, and osteoblasts.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular pathologies. Similarly, polypeptides and antibodies

35

directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, reproductive and skeletal systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, smooth muscle, prostate, and osteoblasts, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 24 are useful for detection and treatment of neoplasias and developmental abnormalities associated with these tissues.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 157 as residues: Asn-21 to Thr-26.

FEATURES OF PROTEIN ENCODED BY GENE NO: 25

The translation product of Gene NO: 25 shares sequence homology with Pregnancy Associated Mouse Protein (PAMP)-1. (See, FEBS Lett 1993 May 17;322(3):219-222). Based on the sequence similarity the translation product of this gene is expected to share certain biological activities with PAMP-1.

Gene NO: 25 is expressed primarily in 12-week-old human embryos and prostate.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, prostate disorders (cancer). Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the prostate, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 25 are useful for the diagnosis and treatment of prostate disorders (such as cancer) and developmental abnormalities and fetal deficiencies. The homology to PAMP-1 indicates that this gene and gene product are useful in detecting pregnancy.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 158 as residues: Pro-23 to Glu-28 and Ser-44 to Gly-55.

FEATURES OF PROTEIN ENCODED BY GENE NO: 26

Gene NO: 26 is expressed primarily in testes and to a lesser extent in epididymis.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, reproductive and endocrine disorders, as well as testicular cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive and endocrine systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes, and epididymis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 26 are useful for the treatment and diagnosis of conditions concerning proper testicular function (e.g., endocrine function, sperm maturation), as well as cancer. Therefore, this gene product is useful in the treatment of male infertility and/or impotence. This gene product is also useful in assays designed to identify binding agents as such agents (antagonists) are useful as male contraceptive agents. Similarly, the protein is believed to be useful in the treatment and/or diagnosis of testicular cancer. The testes are also a site of active gene expression of transcripts that may be expressed, particularly at low levels, in other tissues of the body. Therefore, this gene product may be expressed in other specific tissues or organs where it may play related functional roles in other processes, such as hematopoiesis, inflammation, bone formation, and kidney function, to name a few possible target indications.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 159 as residues: Pro-24 to Gly-33 and Arg-70 to Gly-76.

FEATURES OF PROTEIN ENCODED BY GENE NO: 27

5 The translation product of Gene NO: 27 shares sequence homology with salivary protein precursors which are thought to be important in immune response and production of secreted proteins.

Gene NO: 27 is expressed primarily in salivary gland tissue.

10 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, diseases of the salivary gland. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a
15 number of disorders of the above tissues or cells, particularly of the immune system, digestive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., salivary gland, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder,
20 relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

 The tissue distribution and homology to salivary secreted protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 27 are useful for treatment of immune disorders and diagnostic uses related to secretion of protein in
25 disease states. For example, the gene product can be used as an anti-microbial agent, an ingredient for oral or dental hygiene, treatment of xerostomia, sialorrhea, intervention for inflammation including parotitis, and an indication for tumors in the salivary gland (adenomas, carcinomas).

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
30 160 as residues: Asp-21 to Gly-28, Asp-30 to Glu-43, Glu-49 to Glu-62, and Thr-75 to Pro-83.

FEATURES OF PROTEIN ENCODED BY GENE NO: 28

35 Gene NO: 28 is expressed primarily in human fetal heart tissue and to a lesser extent in olfactory tissue.

 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a

biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune, olfactory and cardiovascular disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, olfactory and vascular systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., olfactory tissue, and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 28 are useful for diagnosis and treatment of immune, olfactory and vascular disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 161 as residues: Cys-33 to Gly-44, Arg-71 to Arg-78, Ser-130 to Gly-142, Lys-150 to Gly-157, and Thr-159 to Asp-177.

FEATURES OF PROTEIN ENCODED BY GENE NO: 29

Gene NO: 29 is expressed primarily in brain and to a lesser degree in activated macrophages, endothelial and smooth muscle cells, and some bone cancers.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of brain and endothelial present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurodegeneration, inflammation and other immune disorders, fibrotic conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification brain, smooth muscle, and endothelium. For a number of disorders of the above tissues or cells, particularly of the brain and endothelium, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., brain, endothelial cells, macrophages, smooth muscle, and bone, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

Tissue distribution suggests polypeptides and polynucleotides corresponding to Gene NO: 29 are useful in study and treatment of neurodegenerative and immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 162 as residues: Asn-18 to Glu-20, Ser-33 to Gln-48, Cys-55 to Ser-56. Pro-67 to Cys-69.

FEATURES OF PROTEIN ENCODED BY GENE NO: 30

Gene NO: 30 is expressed primarily in early stage human brain and to a lesser extent in cord blood, heart, and some tumors.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of developing CNS tissue present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular and neurodegenerative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous and immune systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and heart, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that that polypeptides and polynucleotides corresponding to Gene NO: 30 are useful for the treatment of cancer and of neurodegenerative and cognitive disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 31

Gene NO: 31 is expressed primarily in brain and thymus and to a lesser extent in several other organs and tissues including the hematopoietic system, liver skin and bone

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential

identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly CNS disorders, hematopoietic system disorders, disorders of the endocrine system, bone, and skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g.,

5 hematopoietic cells, brain, thymus, liver, bone, and epidermis, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

10 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 31 are useful for treatment and diagnosis of CNS disorders, hematopoietic system disorders, disorders of the endocrine system, and of bone and skin.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
15 164 as residues: Thr-35 to Arg-40, Pro-55 to His-75, Pro-93 to Ala-98, Ala-111 to Pro-119, and Pro-132 to Glu-138.

FEATURES OF PROTEIN ENCODED BY GENE NO: 32

Gene NO: 32 is expressed primarily in organs and tissue of the nervous system
20 and to a lesser extent in various developing tissues and organs.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, disorders of the central nervous system and disorders of developing and
25 growing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly disorders of the CNS, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., tissue of the nervous
30 system and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

35 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 32 are useful for diagnosis and treatment of disorders of the central nervous system, general neurological diseases and neoplasias.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 165 as residues: Ser-33 to Lys-41 and Glu-86 to Glu-91.

FEATURES OF PROTEIN ENCODED BY GENE NO: 33

- 5 Residues 141-156 in the translation product for Gene NO: 33 as shown in the sequence listing matches phosphopantetheine binding site motifs. Phosphopantetheine (or pantetheine 4' phosphate) is the prosthetic group of acyl carrier proteins (ACP) in some multienzyme complexes where it serves as a 'swinging arm' for the attachment of activated fatty acid and amino-acid groups. Phosphopantetheine is attached to a serine
- 10 residue in these proteins. ACP proteins or domains have been found in various enzyme systems which are listed below. Fatty acid synthetase (FAS), which catalyzes the formation of long-chain fatty acids from acetyl-CoA, malonyl-CoA and NADPH. Bacterial and plant chloroplast FAS are composed of eight separate subunits which correspond to the different enzymatic activities; ACP is one of these polypeptides.
- 15 Fungal FAS consists of two multifunctional proteins, FAS1 and FAS2; the ACP domain is located in the N-terminal section of FAS2. Vertebrate FAS consists of a single multifunctional enzyme; the ACP domain is located between the beta-ketoacyl reductase domain and the C-terminal thioesterase domain. Based on the presence of a phosphopantetheine binding site in the translation product of this gene, it is believed to
- 20 share activities fatty acid synthetase polypeptides. Such activities may be assayed by methods known in the art.

This gene is expressed primarily in developing and rapidly growing tissues like placenta fetal heart and endometrial tumor and to a lesser extent in B and T cell lymphoma tissues

- 25 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and disorders of developing tissues and organs. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
- 30 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic tissues and developing organs and tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., embryonic tissue, endometrium, B-cells, and T-cells, and cancerous and wounded
- 35 tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to

the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 33 are useful for treatment and diagnosis of cancer in the hematopoietic system developing organs and tissues. It may also be useful for induction of cell growth in disorders of the hematopoietic system and other tissue and organs. The homology to fatty acid synthetases indicates that this gene product is useful in the diagnosis and treatment of lipid metabolism disorders such as hyperlipidemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 166 as residues: Arg-27 to Glu-34.

FEATURES OF PROTEIN ENCODED BY GENE NO: 34

Gene NO: 34 is expressed primarily in breast and testes tissues and to a lesser extent in hematopoietic tissues including tonsils, T cells and monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the reproductive organs and systems, including cancer, autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive organs and hematopoietic tissues, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells, T-cells and monocytes, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Nucleic acids comprising sequence of this gene are also useful as chromosome markers since this gene maps to Chr.15, D15S118-D15S123.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 34 are useful for treatment of diseases of the reproductive organs and hematopoietic system including cancer, autoimmune diseases and inflammatory diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 167 as residues: Phe-81 to Lys-86.

FEATURES OF PROTEIN ENCODED BY GENE NO: 35

The translation product of Gene NO: 35 shares sequence similarity with the mouse cytokine-inducible inhibitor of signaling. See, e.g., Nature 1997 Jun 26;387(6636):917-921. Cytokines are secreted proteins that regulate important cellular responses such as proliferation and differentiation. Key events in cytokine signal transduction are well defined: cytokines induce receptor aggregation, leading to activation of members of the JAK family of cytoplasmic tyrosine kinases. In turn, members of the STAT family of transcription factors are phosphorylated, dimerize and increase the transcription of genes with STAT recognition sites in their promoters. Less is known of how cytokine signal transduction is switched off. Expression of the mouse SOCS-1 protein inhibited both interleukin-6- induced receptor phosphorylation and STAT activation. We have also cloned two relatives of SOCS-1, named SOCS-2 and SOCS-3, which together with the previously described CIS form a new family of proteins. Transcription of all four SOCS genes is increased rapidly in response to interleukin-6, in vitro and in vivo, suggesting they may act in a classic negative feedback loop to regulate cytokine signal transduction. The translation product of this gene is believed to have similar biological activities as this family of mouse genes. The biological activity of the translation product of this gene may be assayed by methods shown in Nature 1997 Jun 26;387(6636): 917-921, which is incorporated herein by reference in its entirety.

Gene NO: 35 is expressed primarily in tissues of hematopoietic origin including activated monocytes, neutrophils, activated T-cells and to a lesser extent in breast, adipose tissue and dendritic cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the hematopoietic system including cancer autoimmune diseases and inflammatory diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hematopoietic system expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to cytokine inducible inhibitor of signaling indicates that polypeptides and polynucleotides corresponding to Gene NO: 35 are
5 useful for diagnosis and treatment of diseases of the hematopoietic system including autoimmune diseases, inflammatory diseases, infectious diseases and neoplasia. For example, administration of, or upregulation of this gene could be used to decrease the response of immune-system to lymphokines and cytokines.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
10 168 as residues: Arg-23 to His-30, Ala-35 to Gly-42.

FEATURES OF PROTEIN ENCODED BY GENE NO: 36

Gene NO: 36 is expressed primarily in infant brain and to a lesser extent in osteoclastoma, placenta, and a wide variety of other tissues.

15 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, neurological disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential
20 identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., osteoclastoma, placenta, and tissue of the central nervous system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial
25 fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 36 are useful for diagnosis and treatment of neurologic
30 disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 169 as residues: Gln-31 to Ser-37, Ile-49 to Gly-54, Tyr-57 to Asp-67, Gln-141 to Pro-151, and Val-207 to Thr-219.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 37

Gene NO: 37 is expressed primarily in osteoclastoma stromal cells, dendritic cells, liver, and placenta.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer, wound, pathological conditions. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., stromal cells, dendritic cells, liver, and placenta and, cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 37 are useful for fundamental role in basic growth and development of human.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 170 as residues: Leu-32 to Thr-37 and Arg-48 to Pro-55.

20 FEATURES OF PROTEIN ENCODED BY GENE NO: 38

The translation product of Gene NO: 38 shares sequence homology with a yeast protein, Lpe10p, which may be involved in mRNA processing. (See Accession Nos. 2104457 and 1079682.) It is likely that an upstream signal sequence exists, other than the predicted sequence described in Table 1. Preferred polypeptide fragments comprise the open reading frame upstream from the predicted signal sequence, as well as polynucleotide fragments encoding these polypeptide fragments.

This gene is expressed primarily in skin, and to a lesser extent in embryonic tissues, and fetal liver.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the skin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the skin, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, liver, and embryonic tissues, and cancerous and wounded tissues) or bodily fluids (e.g., serum,

plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 38 are useful for diagnosis and treatment of defects of the skin.

FEATURES OF PROTEIN ENCODED BY GENE NO: 39

- 10 Gene NO: 39 is expressed primarily in Amygdala, activated monocytes, testis, and fetal liver. Moreover, this gene is mapped to chromosome 4. Thus, polynucleotides of the present invention can be used in linkage analysis as markers for chromosome 4.

- Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects of the brain, immune system and testis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, immune system and testis, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., Amygdala, monocytes, testes, and liver and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.
- 15
- 20
- 25

 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 39 are useful for detecting defects of the brain, immune system and testis because of its abundance in these tissues.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 40

 The translation product of Gene NO: 40 shares sequence homology with lymphoma 3-encoded protein (bcl-3) which is thought to contribute to leukemogenesis when abnormally expressed.

- 35 This gene is expressed primarily in Human Neutrophils, and to a lesser extent in Human Osteoclastoma Stromal Cells (unamplified), Hepatocellular Tumor, and Human Neutrophils, (Activated).

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, chronic lymphocytic leukemia. Similarly, polypeptides and antibodies
5 directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell
10 types (e.g., neutrophils, osteoclastoma, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to lymphoma 3-encoded protein (bcl-3)
15 indicates that polypeptides and polynucleotides corresponding to Gene NO: 40 are useful for treatment of lymphoma and related cancers.

FEATURES OF PROTEIN ENCODED BY GENE NO: 41

20 Gene NO: 41 is expressed primarily in ovary tumor, and to a lesser extent in endometrial stromal cells and fetal brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
25 not limited to, ovarian or endometrial cancer. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the female reproductive system and the developing central nervous system, expression of this gene at significantly higher or
30 lower levels may routinely be detected in certain tissues (e.g., ovary, endometrium and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the
35 disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 41 are useful for development of factors involved in ovarian or endometrial and general reproductive organ disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 174 as residues: Glu-22 to Trp-31, Asn-84 to Asp-90, and Ser-144 to Asp-151.

FEATURES OF PROTEIN ENCODED BY GENE NO: 42

The translation product of Gene 42 has sequence identity with a gene designated PTHrP(B). The PTHrP(B) polypeptide inhibits parathyroid hormone related peptide (PTHrP) activity.

This gene is expressed primarily in adult testis, and to a lesser extent in pituitary.

Therefore polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of male reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the male reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., testes, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Furthermore, based in part on sequence identity with PTHrP(B), nucleic acids and polypeptides of the present invention may be used to diagnose or treat such conditions as hypercalcemia, osteoporosis, and disorders related to calcium metabolism.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 42 are useful for treatment of male reproductive disorders, hypercalcemia, osteoporosis, and other disorders related to calcium metabolism.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 175 as residues: Tyr-81 to Met-86, Gly-103 to Ser-108, Glu-127 to Pro-128, Pro-175 to Ser-180, Glu-196 to Lys-203, Pro-235 to Ser-241, and Ala-249 to Ser-264.

FEATURES OF PROTEIN ENCODED BY GENE NO: 43

The translation product of Gene NO: 43 shares sequence homology with brevican, which is thought to be important as a proteoglycan core protein of the

aggrecan/versican family. The translation product of this gene may also contain a hyaluronan (HA)-binding region domain in frame with, but downstream of, the predicted open reading frame (Barta, et al., *Biochem. J.* 292:947-949 (1993)). The HA-binding domain, also termed the link domain, is found in proteins of vertebrates that are involved in the assembly of extracellular matrix, cell adhesion, and migration. It is about 100 amino acids in length. The structure has been shown to consist of two alpha helices and two antiparallel beta sheets arranged around a large hydrophobic core similar to that of C-type lectin. This domain typically contains four conserved cysteines involved in two disulfide bonds.

10 This gene is expressed primarily in early stage human brain and to a lesser extent in frontal cortex and epileptic tissues.

 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of disorders associated with, or observed during, neuronal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of neuronal and associated tissues and cell types. For a number of disorders of the above tissues or cells, particularly for those of the nervous system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

25 The tissue distribution and homology to brevican indicates that polypeptides and polynucleotides corresponding to Gene NO: 43 are useful for neuronal regulation and signaling. The uses include directing or inhibiting axonal growth for the treatment of neuro-fibromatosis and in detection of glioses.

 Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 176 as residues: Asp-28 to Arg-33 and Arg-126 to Arg-131.

FEATURES OF PROTEIN ENCODED BY GENE NO: 44

 Gene NO: 44 is the human homolog of Notch-2 (Accession No. 477495) and mouse EGF repeat transmembrane protein (Accession No. 1336628), both genes are important in differentiation and development of an organism. The EGF repeat transmembrane protein is regulated by insulin like growth factor Type I receptor. These proteins are involved in cell-cell signaling and cell fate determination. Based on

homology, it is likely that this gene products also involved in cell differentiation and development. Although the predicted signal sequence is indicated in Table 1, it is likely that a second signal sequence is located further upstream. Moreover, further translated coding regions are likely found downstream from the disclosed sequence, which can easily be obtained using standard molecular biology techniques. A frameshift occurs somewhere around nucleotide 714, causing a frame shift in amino acid sequence from frame +2 to frame +3. However, using the homology of Notch-2 and EGF repeat transmembrane protein, the complete open reading frame can be elucidated. Preferred polynucleotide fragments comprise nucleotides 146-715, 281-715, and 714-965. Other preferred polypeptide fragments comprise the following EGF-like motifs:

5
10
CRCASGFTGEDC (SEQ ID NO:260), CTCQVGFTGKEC (SEQ ID NO:261),
CLNLPGSYQCQC (SEQ ID NO:262), CKCLTGFTGQKC (SEQ ID NO:263), and
CQCLQGFTGQYC (SEQ ID NO:264).

Gene NO: 44 is expressed primarily in placenta and to a lesser extent in stromal and immune cells.

15

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hemophilia and other blood disorders, central nervous system disorders, muscle disorders, and any other disorder resulting from abnormal development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune, hematopoietic and vascular systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, stromal and immune cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

20
25
30

The tissue distribution and homology to Notch-2 indicates that polypeptides and polynucleotides corresponding to Gene NO: 44 are useful for diagnosing and treating disorders relating to abnormal regulation of cell fate, induction, and differentiation of cells (e.g., cancer), epidermal growth factors, axonal pathfinding, and hematopoiesis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 177 as residues: Gln-27 to Tyr-32, His-45 to Glu-55, Tyr-61 to Gly-77, Glu-99 to Ser-106, Ser-125 to Cys-131, and Thr-138 to Trp-144.

35

FEATURES OF PROTEIN ENCODED BY GENE NO: 45

The translation product of this gene shares sequence homology with Laminin A which is thought to be important in the binding of epithelial cells to basement membrane and is associated with tumor invasion. Moreover, the translated protein is homologous to the *Drosophila* LAMA gene (Accession No. 1314864), a gene expressed in the first optic ganglion of *Drosophila*. Thus, it is likely that the gene product from this gene is involved in the development of the eye. Nucleotide fragments comprising nucleotides 822-1223, 212-475, 510-731, and 1677-1754 are preferred. Also preferred are the polypeptide fragments encoded by these polynucleotide fragments. It is likely that a frame shift occurs somewhere between nucleotides 475 to 510, shifting the open reading frame from +2 to +3. However, the open reading frame can be clarified using known molecular biology techniques.

This gene is expressed primarily in human testes tumor and to a lesser extent in placenta and activated monocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, invasive cancers or tumors of the epithelium, as well as disorders relating to eye development. Similarly, polypeptides and antibodies directed to these polypeptides are useful as immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of neoplastic conditions, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., testes, placenta, and monocytes and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to Laminin A indicates that polypeptides and polynucleotides corresponding to Gene NO: 45 are useful for study and diagnosis of malignant or benign tumors, fibrotic disorders, and eye disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 178 as residues: Met-1 to Gly-8, Glu-32 to Ala-37, Met-113 to Asn-119, and Glu-139 to Gln-153.

FEATURES OF PROTEIN ENCODED BY GENE NO: 46

The translation product of Gene NO: 46 is novel and shares sequence homology with the product of the *Drosophila* tissue polarity gene frizzled. In vertebrates, it appears that there is a family of proteins that represent frizzled gene homologs. (See, 5 e.g., Accession Nos. 1946343 and AFO17989.) The *Drosophila* frizzled protein is thought to transmit polarity signals across the plasma membrane of epidermal cells. The structure of frizzled proteins suggest that they may function as a G-protein-coupled receptor. The frizzled proteins are thought to represent receptors for Wnt gene products - secreted proteins that control tissue differentiation and the development of embryonic 10 and adult structures. Inappropriate expression of Wnts has also been demonstrated to contribute to tumor formation. Moreover, mammalian secreted frizzled related proteins are thought to regulate apoptosis. (See Accession No. AFO17989.) The human homolog has also been recently cloned by other groups. (See Accession No. H2415415.) Thus, the protein encoded by this gene plays a role in mediating tissue 15 differentiation, proliferation, tumorigenesis and apoptosis. Preferred polypeptide fragments lack the signal sequence as described in Table 1, as well as N-terminal and C-terminal deletions. Preferred polynucleotide fragments encode these polypeptide fragments.

Gene NO: 46 is expressed primarily in fetal tissues - particularly fetal lung - and 20 adult cancers, most notably pancreas tumor and Hodgkin's lymphoma. Together, this distribution is consistent with expression in tissues undergoing active proliferation. The gene is also expressed to a lesser extent in other organs, including stomach, prostate, and thymus.

Therefore, polynucleotides and polypeptides of the invention are useful as 25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer (particularly pancreatic cancer and/or Hodgkin's lymphoma), as well as other forms of aberrant cell proliferation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for 30 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system and hyperproliferative disorders, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., fetal tissue, pancreas, and tissue of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, 35 urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e.,

the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to frizzled indicates that polypeptides and polynucleotides corresponding to Gene NO: 46 are useful for influencing cell proliferation, differentiation, and apoptosis. The full-length protein or a truncated domain could potentially bind to and regulate the function of specific factors, such as Wnt proteins or other apoptotic genes, and thereby inhibit uncontrolled cellular proliferation. Expression of this protein within a cancer - such as via gene therapy or systemic administration - could effect a switch from proliferation to differentiation, thereby arresting the progression of the cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 179 as residues: Pro-31 to Arg-37.

FEATURES OF PROTEIN ENCODED BY GENE NO: 47

The translation product of Gene NO: 47 shares sequence homology with members of the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes. These ribonuclease proteins are found predominantly in fungi, plants, and bacteria and have been implicated in a number of functions, including phosphate-starvation response, self-incompatibility, and responses to wounding. A second group has recently cloned this same gene, calling it a ribonuclease 6 precursor. (See Accession No. 2209029.) This group also mapped the gene to chromosome 6, thus, the polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 6.

Gene NO: 47 is expressed primarily in hematopoietic cells and tissues, including macrophages, eosinophils, CD34 positive cells, T-cells, and spleen. It is also expressed to a lesser extent in brain and spinal cord.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a hematopoietic origin, graft rejection, wounding, inflammation, and allergy. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., hematopoietic cells, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder.

relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the Rh/T2/S-glycoprotein family of ribonuclease-encoding genes indicates that polypeptides and polynucleotides
5 corresponding to Gene NO: 47 are useful as a cytotoxin that could be directed against specific cell types (e.g. cancer cells; HIV- infected cells), and that would be well tolerated by the human immune system.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 180 as residues: Ala-24 to Asp-30, Ile-51 to Tyr-61, Pro-69 to Ser-78, Pro-105 to Phe-
10 110, Asn-129 to Phe-135, Pro-187 to Glu-192, Lys-205 to Gln-224, and Pro-250 to His-256.

FEATURES OF PROTEIN ENCODED BY GENE NO: 48

The translation product of Gene NO: 48 shares sequence homology with
15 dolichyl-phosphate glucosyltransferase, a transmembrane-bound enzyme of the endoplasmic reticulum which is thought to be important in N-linked glycosylation, by catalyzing the transfer of glucose from UDP-glucose to dolichyl phosphate. (See Accession No. 535141.) Based on homology, it is likely that this gene product also play a role similar in humans. Preferred polynucleotide fragments comprise nucleotides
20 132-959. Also preferred are the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 48 is expressed primarily in endothelial cells and to a lesser extent in hematopoietic cells and brain.

Therefore, polynucleotides and polypeptides of the invention are useful as
25 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, defects in proper N-linked glycosylation of proteins, such as Wiskott-Aldrich syndrome; tumors of an endothelial cell origin. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes
30 for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the vascular and hematopoietic systems, as well as brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., endothelial cells, hematopoietic cells, and brain, and cancerous and wounded tissues) or bodily fluids
35 (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene

expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to dolichyl-phosphate glucosyltransferase indicates that polypeptides and polynucleotides corresponding to Gene NO: 48 are
5 useful in diagnosing and treating defects in N-linked glycosylation pathways that contribute to disease conditions and/or pathologies.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 181 as residues: Lys-50 to Thr-55, Ser-73 to Arg-79, Glu-92 to Pro-99, Asp-110 to Ser-117, Gln-125 to Lys-131, Gly-179 to Asn-188, Ile-231 to Cys-236, and Glu-318
10 to Asn-324.

FEATURES OF PROTEIN ENCODED BY GENE NO: 49

Gene NO: 49 is expressed primarily in brain, most notably in the hypothalamus and amygdala. This gene is also mapped to chromosome X, and therefore, can be used
15 in linkage analysis as a marker for chromosome X.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, tumors of a brain origin; neurodegenerative disorders, and sex-linked
20 disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., brain and cancerous and wounded tissues) or bodily
25 fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides
30 corresponding to Gene NO: 49 are useful for the diagnosis of tumors of a brain origin, and the treatment of neurodegenerative disorders, such as Parkinson's disease, and sex-linked disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 50

35 The translation product Gene NO: 50 shares sequence homology with canine phospholemman, a major plasma membrane substrate for cAMP-dependent protein kinases A and C. (See Accession No. M63934; see also Accession No. A40533.) In

fact, a group also recently cloned the human phospholemman gene, and mapped this gene to chromosome 19. (See Accession No.1916010.) Phospholemman is a type I integral membrane protein that gets phosphorylated in response to specific extracellular stimuli such as insulin and adrenalin. Phospholemman forms ion channels in the cell membrane and appears to regulate taurine transport, suggesting an involvement in cell volume regulation. It has been proposed that phospholemman is a member of a superfamily of membrane proteins, characterized by single transmembrane domains, which function in transmembrane ion flux. They are capable of linking signal transduction to the regulation of such cellular processes as the control of cell volume.

Gene No 50 is expressed primarily in fetal liver and to a lesser extent in adult brain and kidney, as well as other organs.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, insulin and/or adrenalin defects; diabetes; aberrant ion channel signaling; defective taurine transport; and defects in cell volume regulation. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the brain and/or immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, brain, and kidney, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to phospholemman indicates that polypeptides and polynucleotides corresponding to Gene NO: 50 are useful for treatment of disorders involving the transport of ions and small molecules, in particular taurine. It could also be beneficial for control of pathologies or diseases wherein aberrancies in the control of cell volume are a distinguishing feature, due to the predicted role for phospholemman in the normal control of cell volume. It also may play a role in disorders involving abnormal circulating levels of insulin and/or adrenalin - along with other active secreted molecules - as revealed by its phosphorylation upon stimulation with insulin or adrenalin.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 183 as residues: Ala-20 to Gln-34, Arg-58 to Thr-79, and Leu-87 to Arg-92.

FEATURES OF PROTEIN ENCODED BY GENE NO: 52

Gene NO: 52 is expressed primarily in metastatic melanoma and to a lesser extent in infant brain.

- 5 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and cancer metastasis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for
- 10 differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., epidermis, and brain, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a
- 15 disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 52 are useful for diagnosis and treatment of melanoma.

20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 53**

The translation product of Gene NO: 53 shares sequence homology with mucin which is thought to be important cell surface molecule. It also exhibits sequence identity with a calcium channel blocker of *Agelenopsis aperta*. In particular, with those calcium channel blockers which affect neuronal and muscle cells.

- 25 Gene NO: 53 is expressed primarily in prostate, endothelial cells, smooth muscle and fetal tissues and to a lesser extent in T cells and placenta.

- Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are
- 30 not limited to, prostate cancer, immune disorders, angina, hypertension, cardiomyopathies, supraventricular arrhythmia, oesophageal achalasia, premature labour, and Raynaud's disease. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells,
- 35 particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues or cell types (e.g., prostate, and tissue and cells of the immune system, and cancerous and wounded tissues) or

bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 5 The tissue distribution and homology to mucin indicates that polypeptides and polynucleotides corresponding to Gene NO: 53 are useful as a surface antigen for diagnosis of diseases such as prostate cancer and as tumor vaccine.

FEATURES OF PROTEIN ENCODED BY GENE NO: 54

- 10 Gene NO: 54 encodes a polypeptide which exhibits sequence identity with the rab receptor and VAMP-2 receptor proteins. (Martincic, et al., J. Biol. Chem. 272 (1997).)

Gene NO: 54 is expressed primarily in placenta, fetal liver, osteoclastoma and smooth muscle and to a lesser extent in T cell, fetal lung and colon cancer.

- 15 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancers, osteoporosis and immuno-related diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing
- 20 immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, hematopoiesis system and bone system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, liver, osteoclastoma, smooth muscle, T-cells, and lung, and colon, and
- 25 cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

- 30 The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 54 are useful for treating cancer, osteoporosis and immuno-disorders.

- Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 187 as residues: Pro-16 to Phe-21, Pro-24 to Arg-35, Arg-92 to Pro-98, Asn-143 to
- 35 Lys-151, and Leu-169 to Ile-176.

FEATURES OF PROTEIN ENCODED BY GENE NO: 55

Gene NO: 55 encodes a protein having sequence identity to the rat galanin receptor GALR2.

Gene NO: 55 is expressed primarily in ovarian cancer.

- 5 Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of ovarian cancer. Similarly, polypeptides and antibodies directed to those polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders
- 10 of the above tissues or cells, particularly of the immune system and reproductive system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., ovary, and tissues and cells of the immune system, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample
- 15 taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. GALR2 antagonists can be used to treat obesity, bulimia, or Alzheimer's disease, while GALR2 agonists can be used to treat anorexia or pain, or to decrease nociception (claimed). Agonists and antagonists can also be used to
- 20 treat numerous other disorders, including cognitive disorders, sensory disorders, motion sickness, convulsion/epilepsy, hypertension, diabetes, glaucoma, reproductive disorders, gastric and intestinal ulcers, inflammation, immune disorders, and anxiety.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 55 are useful for diagnosis and treatment of ovarian cancer.

25

FEATURES OF PROTEIN ENCODED BY GENE NO: 56

- As indicated in Table 1, the predicted signal sequence of Gene NO: 56 relates to an open reading frame that is homologous to the mouse major histocompatibility locus class III. (See Accession No. 2564953.) Any frame shift mutations that alter the correct
- 30 open reading frame can easily be clarified using known molecular biology techniques. Moreover, in the opposite orientation, a second translated product is disclosed. This second translation product of this contig is identical in sequence to intracellular protein lysophosphatidic acid acyltransferase. The nucleotide and amino acid sequences of this translated product have since been published by Stamps and colleagues (Biochem. J.
- 35 326 (Pt 2), 455-461 (1997)), West and coworkers (DNA Cell Biol. 6, 691-701 (1997)), Rowan (GenBank Accession No. U89336), and Soyombo and Hofmann (GenBank Accession No. AF020544). This gene is thought to enhance cytokine

signaling response in cells. It is likely that a signal peptide is located upstream from this translated product. Preferred polypeptide fragments comprise the amino acid sequence: GLACWLAGVIFIDRKRTGDAISVMSEVAQTLLTQDVXVWVFPEGTRNHNGSML PFKRGAFHLAVQAQVPIVPIVMSSYQDFYCKKERRFTSGQCQVRVLPVPTEGL
 5 TPDVPALADRVRHSMLHCF (SEQ ID NO: 265);
 PSAKYFFKMAFYNGWILFLAVLAIPVCAVRGRNVENMKILRLMLLHIKYLYGI
 RVEVRGAHHFPPSQPYVVVSNHQSSDLLGMMEVLPGRVCVPIAKR (SEQ ID
 NO:266); TVFREISTD (SEQ ID NO:267); or LWAGSAGWPAG (SEQ ID NO: 268).
 Also provided are polynucleotide fragments encoding these polypeptide fragments.

10 Gene NO: 56 is expressed primarily in infant adrenal gland, hypothalamus, 7
 week old embryonic tissue, fetal lung, osteoclastoma stromal cells, and to a lesser
 extent in a large number of additional tissues.

Therefore, polynucleotides or polypeptides of the invention are useful as
 reagents for differential identification of the tissue(s) or cell type(s) present in a
 15 biological sample and for diagnosis of developmental disorders and osteoclastoma.
 Similarly, polypeptides and antibodies directed to these polypeptides are useful in
 providing immunological probes for differential identification of the tissue(s) or cell
 type(s) in which it is highly expressed. For a number of disorders of the above tissues
 or cells, particularly during development or of the nervous or bone systems, expression
 20 of this gene at significantly higher or lower levels may routinely be detected in certain
 tissues and cell types (e.g., adrenal, embryonic tissue, lung, and osteoclastoma stromal
 cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine,
 synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell
 sample taken from an individual having such a disorder, relative to the standard gene
 25 expression level, i.e., the expression level in healthy tissue or bodily fluid from an
 individual not having the disorder. Further, expression of this protein can be used to
 alter the fatty acid composition of a given cell or membrane type.

The tissue distribution indicates that polypeptides and polynucleotides
 corresponding to Gene NO: 56 are useful for diagnosis and treatment of osteoclastoma
 30 and other bone and non-bone-related cancers, as well as for the diagnosis and treatment
 of developmental disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO:
 189 as residues: Gly-29 to Gly-36 and Tyr-49 to Tyr-58.

35 **FEATURES OF PROTEIN ENCODED BY GENE NO: 57**

The translation product of Gene NO: 57 shares sequence homology with
 longevity-assurance protein-1. (See Accession No. g1123105.) Preferred

polynucleotide fragments comprise nucleotides 6-125 and 118-432, as well as the polypeptides encoded by these polynucleotides. It is likely that a second signal sequence exists upstream from the predicted signal sequence in Table 1. Moreover, a frame shift likely occurs between nucleotides 118-125, which can be elucidated using standard molecular biology techniques.

Gene NO: 57 is expressed primarily in fetal liver, kidney, brain, thymus, and bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immunological diseases and hyperproliferative disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the fetal liver, kidney, brain, thymus, and bone marrow expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, kidney, brain, thymus, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to longevity-assurance protein suggest that Gene NO: 57 encodes a protein useful in increasing life span and in replacement therapy for those suffering from immune system disorders or hyperproliferative disorders caused by underexpression or overexpression of this gene.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 190 as residues: Val-29 to Arg-46 and Gly-50 to Gly-56.

FEATURES OF PROTEIN ENCODED BY GENE NO: 58

Domains of the Gene NO: 58 product are homologous to porcine surfactant protein-A receptor. (See Accession No. B48516.) The bovine gene binds surfactant protein-A receptor, modulating the secretion of alveolar surfactant. Based on this homology, the gene product encoded by this gene will likely have activity similar to the porcine gene. Preferred polynucleotide fragments comprise nucleotides 887-1039, as well as the polypeptide fragments encoded by this nucleotide fragment.

Gene NO: 58 is expressed primarily in brain and to a lesser extent in endothelial cells.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the central nervous system including dementia, stroke, neurological disorders, respiratory distress, and diseases affecting the endothelium including inflammatory diseases, restenosis, and vascular diseases. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the placenta, liver, endothelial cells, prostate, thymus, and lung, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, and endothelial cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology indicates that polypeptides and polynucleotides corresponding to Gene NO: 58 are useful for the diagnosis and /or treatment of diseases on the central nervous system, such as a factor that promote neuronal survival or protection, in the treatment of inflammatory disorders of the endothelium, or in disorders of the lung. In addition this protein may inhibit or promote angiogenesis and therefore is useful in the treatment of vascular disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 191 as residues: His-66 to Pro-80, Gly-139 to Ser-146 and Ser-262 to Pro-267.

FEATURES OF PROTEIN ENCODED BY GENE NO: 59

The translation product of Gene NO: 59 is homologous to the rat hypertension-induced protein which is thought to be important in hypertension, and found expressed mainly in kidneys. (See Accession No. B61209.) Thus, it is likely that this gene product is involved in hypertension in humans. Preferred polypeptide fragments comprise the short chain dehydrogenase/reductase motif SILGIISVPLSIGYCASKHALRGFFNGLR (SEQ ID NO:269), as well as polynucleotides encoding this polypeptide fragment. Also preferred are polynucleotide fragments of 337-639, as well as the polypeptide fragments encoded by this polynucleotide fragment.

Gene NO: 59 is expressed primarily in liver, spleen, lung, brain, and prostate.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cardiovascular, immunological, and renal disorders. Similarly,

5 polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the cardiovascular, renal, and immune, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., liver, spleen, lung, brain, and prostate, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

15 The tissue distribution and homology to hypertension-induced protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 59 are useful for treating hypertension.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 192 as residues: Gln-40 to Glu-45, Glu-96 to Glu-102, Asn-256 to Thr-266, and Asp-20 308 to Asp-317.

FEATURES OF PROTEIN ENCODED BY GENE NO: 60

Gene NO: 60 is expressed primarily in activated T-cell and Jurkat cell and to a lesser extent in apoptotic T-cell and CD34+ cell. It is likely that alternative open reading frames provide the full length amino acid sequence, which can be verified using standard molecular biology techniques.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, T lymphocyte related diseases or hematopoiesis. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard

gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 60 are useful for diagnosis or treatment of immune system disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 61

The translation product of Gene NO: 61, a vacuolar proton-ATPase, shares sequence homology with a *Caenorhabditis elegans* protein which is thought to be important in development. This protein may be a human secretory homologue that may also influence embryo development. Ludwig, J., also recently cloned this gene from chromaffin granules. (See, Accession No. 2584788.) Although Table 1 indicates the predicted signal peptide sequence, the translated product of this gene may in fact start with the upstream methionine, beginning with the amino acid sequence MAYHGLTV (SEQ ID NO:270). Thus, polypeptides comprising this upstream sequence, as well as N-terminus deletions, are also contemplated in the present invention.

Gene NO: 61 is expressed primarily in human placenta, liver, and Hodgkin's Lymphoma and to a lesser extent in bone marrow. Modest levels of expression were also observed in dendritic cells.

Therefore, polynucleotides and polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, hyperproliferative disorders, defects in embryonic development, and diseases or disorders caused by defects in chromaffin granules. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., placenta, liver, lymph tissue, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* indicates that polypeptides and polynucleotides corresponding to Gene NO: 61 are useful for diagnostic or therapeutic modalities for hyperproliferative disorders, embryonic development disorders, and chromaffin granules disorders.

FEATURES OF PROTEIN ENCODED BY GENE NO: 62

The translation product of Gene NO: 62 shares sequence homology with the murine LAG3 gene which is thought to be important in the mediation of natural killer cell (NK cell) activity as previously determined by experiments in mice containing null mutations of LAG3. The similarity of this gene to the CD4 receptor may imply that the gene product may be a secreted, soluble receptor and immune mediator.

Gene NO: 62 is expressed primarily in human fetal heart, meningima, and to a lesser extent in tonsils. This gene also is expressed in the breast cancer cell line MDA 36.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, lymphomas, leukemias, breast cancer and any immune system dysfunction, including those dysfunctions which involve natural killer cell activities. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system or breast cancer, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., heart, meningima, and tonsils and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the LAG3 gene (murine) indicates that the polynucleotides and polypeptides corresponding to Gene NO: 62 are useful for diagnostic and/or therapeutic modalities directed at abnormalities or disease states involving defective immune systems, preferably involving natural killer cell activity, as well as breast cancer.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 195 as residues: Pro-10 to Trp-17, Cys-58 to Pro-67, Thr-76 to Glu-85, and Arg-93 to Asn-101.

35 FEATURES OF PROTEIN ENCODED BY GENE NO: 63

The translation product of Gene NO: 63 shares sequence homology with a *Caenorhabditis elegans* alpha-collagen gene (Clg), which is thought to be important in

organism development, as well as other collagen genes. Thus, based on sequence homology, polypeptides of this gene are expected to have activity similar to collagen, including involvement in organ development.

Gene NO: 63 is expressed primarily in human B-Cell Lymphoma, and to a lesser extent in human pituitary tissue. This gene has also demonstrated expression in keratinocytes.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, B-Cell Lymphoma, other lymphomas, leukemias, and other cancers, as well as disorders related to development. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., tissue and/or cells of the immune system, and pituitary, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to *Caenorhabditis elegans* alpha-collagen gene indicates that polypeptides and polynucleotides corresponding to Gene NO: 63 are useful for development of diagnostic and/or therapeutic modalities directed at the detection and/or treatment of cancer, specifically B-Cell Lymphomas, leukemias, or diseases related to development.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 196 as residues: Thr-22 to Arg-27 and Ser-29 to Thr-39.

FEATURES OF PROTEIN ENCODED BY GENE NO: 64

The translation product of Gene NO: 64 shares sequence homology with human extracellular molecule olfactomedin, which is thought to be important in the maintenance, growth, or differentiation of chemosensory cilia on the apical dendrites of olfactory neurons. Based on this sequence homology, it is likely that polypeptides of this gene have activity similar to the olfactomedin, particularly the differentiation or proliferation of neurons.

Gene NO: 64 is expressed primarily in fetal lung tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, diseases of the lung as well as neural development, particularly of the lung. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the pulmonary system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., lungs and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the olfactomedin family indicates that polypeptides and polynucleotides corresponding to Gene NO: 64 are useful for the development of diagnostic and/or therapeutic modalities directed at detection and/or treatment of pulmonary disease states, e.g., cystic fibrosis.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 197 as residues: Gly-17 to Gln-23, Gln-45 to Arg-50, Arg-56 to Lys-61, Glu-70 to Leu-76, Asp-88 to Glu-93, Pro-117 to Met-131, Asp-161 to Glu-167, Arg-224 to Asn-237, Asp-302 to Trp-312, Pro-315 to Asn-320, and Thr-337 to Ser-341.

FEATURES OF PROTEIN ENCODED BY GENE NO: 65

The translation product of Gene NO: 65 shares sequence homology with *Saccharomyces cerevisiae* hypothetical protein YKL166 (Accession No. gi/687880) which is thought to be important in secretory and/or vesicular transport mechanisms. Based on this homology, it is likely that the gene product would have similar activity to YKL166, particularly secretory or transport mechanisms. Preferred polypeptide fragments of this gene include those fragments starting with the amino acid sequence ISAARV (SEQ ID NO:271). Other polypeptide fragments include the former fragment, which ends with the amino acid sequence PDVSEFMTRLF (SEQ ID NO:272). Further preferred fragments include those polypeptide fragments comprising the amino acid sequence FDPVRVDITSKGKMRAR (SEQ ID NO:273). Also preferred are polypeptide fragments having exogenous signal sequences fused to the polypeptide.

Gene No 65 is expressed primarily in placenta, testis, osteoclastoma and to a lesser extent in adrenal gland.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or diseases involving defects in protein secretion. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the reproductive system, cartilage and bone, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., placenta, testis, adrenal gland, and osteoclastoma, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to the yeast YKL1GG protein indicates that polypeptides and polynucleotides corresponding to Gene NO: 65 are useful for the development of therapeutic and/or diagnostic modalities targeted at cancer or secretory anomalies, such as genetically caused secretory diseases.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 198 as residues: Ser-18 to Ser-29 and Lys-53 to Arg-74.

FEATURES OF PROTEIN ENCODED BY GENE NO: 66

The translation product of Gene NO: 66 shares sequence homology with the human papilloma virus (HPV) E5 ORF region which is thought to be important as a secreted growth factor. Although this is described as a viral gene product, it is believed to have several cellular secretory homologues. Therefore, based on the sequence similarity between the HPV E5 ORF and the translated product of this gene, this gene product is likely to have activity similar to HPV E5 ORF.

Gene NO: 66 is expressed primarily in activated T-Cells, monocytes, cerebellum and to a lesser extent in infant brain.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, cancer and/or human papilloma virus infection. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of

this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., brain, lymph tissue, monocytes, and T-cells and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder. Moreover, polynucleotides of this gene have been mapped to chromosome 1. Therefore, polynucleotides of the present invention can be used in linkage analysis as a marker for chromosome 1.

The tissue distribution and homology to human papilloma virus E5 region indicates that polypeptides and polynucleotides corresponding to Gene NO: 66 are useful for development of diagnostic and/or therapeutic modalities directed at the diagnosis and/or treatment of cancer and/or human papilloma virus infection (HPV).

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 199 as residues: Asn-31 to Arg-36 and Leu-102 to Ser-112.

FEATURES OF PROTEIN ENCODED BY GENE NO: 67

The translation product of Gene NO: 67 shares sequence homology with the 8hs20 protein precursor [*Mus musculus*] which is thought to be important in B-Cell mu chain assembly. (See, Accession No. PID/d1002996; Shiraswa, T., EMBO. J. 12(5):1827-1834 (1993).) A polypeptide fragment starting at amino acid 53 is preferred, as well as 1-20 amino acid N-terminus and/or C-terminus deletions. Based on the sequence similarity between 8hs20 protein and the translation product of this gene, the two polypeptides are expected to share certain biological activities, particularly immunologic activities.

Gene NO: 67 is expressed primarily in human B-cells and to a lesser extent in Hodgkin's Lymphoma. It is also likely that the polypeptide will be expressed in B-cell specific cells, bone marrow, and spleen, as is observed with 8hs20.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, Hodgkin's Lymphoma, Common Variable Immunodeficiency, and/or other B-cell lymphomas. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., bone

marrow, spleen, lymph tissue, and B-cells, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution and homology to 8hs20 protein precursor [*Mus musculus*], indicates that polypeptides and polynucleotides corresponding to Gene NO: 67 are useful for therapeutic and/or diagnostic purposes, targeting Hodgkin's Lymphoma, B-cell lymphomas, Common Variable Immunodeficiency, or other immune disorders.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 200 as residues: Asp-51 to Trp-56, Arg-72 to Asp-85, and Gln-106 to Asp-112.

FEATURES OF PROTEIN ENCODED BY GENE NO: 68

Gene NO: 68 is expressed primarily in fetal liver/spleen, rhabdomyosarcoma, and to a lesser extent in 9 week-old early stage human embryo and bone marrow.

Therefore, polynucleotides or polypeptides of the invention are useful as reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, rhabdomyosarcoma and other cancers, hematopoietic disorders, and immune dysfunction. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune system, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues (e.g., embryonic tissue, striated muscle, liver, spleen, and bone marrow, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that the protein product of Gene NO: 68 is useful for diagnostic and/or therapeutic purposes directed to cancer, preferably rhabdomyosarcoma. Enhanced expression of this gene in fetal liver, spleen, and bone marrow indicates that this gene plays an active role in hematopoiesis. Polypeptides or polynucleotides of the present invention may therefore help modulate survival, proliferation, and/or differentiation of various hematopoietic lineages, including the hematopoietic stem cell. Thus, polynucleotides or polypeptides can be used treat

various hematopoietic disorders and influence the development and differentiation of blood cell lineages, including hematopoietic stem cell expansion. The polypeptide does contain a thioredoxin family active site at amino acids 64-82. Polypeptides comprising this thioredoxin active site are contemplated.

5

FEATURES OF PROTEIN ENCODED BY GENE NO: 69

Gene NO: 69 is expressed primarily in liver and kidney and to a lesser extent in macrophages, uterus, placenta, and testes.

Therefore, polynucleotides or polypeptides of the invention are useful as
10 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, renal disorders, neoplasms (e.g., soft tissue cancer, hepatocellular tumors), immune disorders, endocrine imbalances, and reproductive disorders. Similarly, polypeptides and antibodies directed to these polypeptides are useful in
15 providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the hepatic, urogenital, immune, and reproductive systems, expression of this gene at significantly higher or lower levels may routinely be detected in certain tissues and cell types (e.g., liver, kidney, uterus, placenta, testes, and macrophages and cancerous and
20 wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder. relative to the standard gene expression level, i.e., the expression level in healthy tissue or bodily fluid from an individual not having the disorder.

The tissue distribution indicates that polypeptides and polynucleotides
25 corresponding to Gene NO: 69 are useful for diagnosis and treatment of disorders in the hepatic, urogenital, immune, and reproductive systems.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 202 as residues: Arg-41 to Ser-50, Glu-138 to Asn-148, Ser-155 to Arg-172, Pro-219 to Glu-228.

30

FEATURES OF PROTEIN ENCODED BY GENE NO: 70

Gene NO: 70 is expressed primarily in the immune system, including macrophages, T-cells, and dendritic cells and to a lesser extent in fetal tissue.

Therefore, polynucleotides or polypeptides of the invention are useful as
35 reagents for differential identification of the tissue(s) or cell type(s) present in a biological sample and for diagnosis of diseases and conditions which include, but are not limited to, immune disorders, inflammatory diseases, lymph node disorders, fetal

development, and cancers. Similarly, polypeptides and antibodies directed to these polypeptides are useful in providing immunological probes for differential identification of the tissue(s) or cell type(s). For a number of disorders of the above tissues or cells, particularly of the immune and hematopoietic systems expression of this gene at

5 significantly higher or lower levels may routinely be detected in certain tissues and certain cell types (e.g., macrophages, T-cells, dendritic cells, and fetal tissue, and cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) or another tissue or cell sample taken from an individual having such a disorder, relative to the standard gene expression level, i.e., the expression level

10 in healthy tissue or bodily fluid from an individual not having the disorder. There is some evidence that the polynucleotide is mapped to chromosome 19. Thus, the polynucleotide can be a marker for genetic analysis for chromosome 19.

The tissue distribution indicates that polypeptides and polynucleotides corresponding to Gene NO: 70 are useful for treatment, prophylaxis, and diagnosis of

15 immune and autoimmune diseases, such as lupus, transplant rejection, allergic reactions, arthritis, asthma, immunodeficiency diseases, leukemia, and AIDS. The polypeptides or polynucleotides of the present invention are also useful in the treatment, prophylaxis, and detection of thymus disorders, such as Graves Disease, lymphocytic thyroiditis, hyperthyroidism, and hypothyroidism. The expression observed

20 predominantly in hematopoietic cells also indicates that the polynucleotides or polypeptides are important in treating and/or detecting hematopoietic disorders, such as graft versus host reaction, graft versus host disease, transplant rejection, myelogenous leukemia, bone marrow fibrosis, and myeloproliferative disease. The polypeptides or polynucleotides are also useful to enhance or protect proliferation, differentiation, and

25 functional activation of hematopoietic progenitor cells (e.g., bone marrow cells), useful in treating cancer patients undergoing chemotherapy or patients undergoing bone marrow transplantation. The polypeptides or polynucleotides are also useful to increase the proliferation of peripheral blood leukocytes, which can be used in the combat of a range of hematopoietic disorders, including immunodeficiency diseases, leukemia, and

30 septicemia.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO: 203 as residues: Thr-21 to Ser-27, Pro-33 to Ser-38, and Arg-73 to Lys-84.

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
1	HGCMID20	97901 02/26/97 209047 05/15/97	pSport1	11	1739	25	1658	54	54	134	1	28	29	466
2	HLDBG33	97898 02/26/97 209044 05/15/97	pCMVSPORT 3.0	12	844	1	844	39	39	135	1	28	29	221
2	HLDBG33	97898 02/26/97 209044 05/15/97	pCMVSPORT 3.0	81	795	1	434	10	10	204	1	29	30	34
3	HTGEW86	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	13	776	134	676	173	173	136	1	35	36	155
4	HKCSR70	97900 02/26/97 209046 05/15/97	pBluescript	14	1376	727	1343	202	202	137	1	20	21	232
4	HKCSR70	97900 02/26/97 209046 05/15/97	pBluescript	82	1324	741	1309		861	205	1	31	32	42

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
4	HETBI87	209010 04/28/97 209085 05/29/97	Uni-ZAP XR	83	1494	1	1484	51	51	206	1	34	35	84
5	HTEAU17	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	15	502	1	502	143	143	138	1	33	34	60
6	HBMCY91	97897 02/26/97 209043 05/15/97	pBluescript	16	425	1	425	56	56	139	1	17	18	72
7	HSSGE07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	17	1316	1	1298	45	45	140	1	26	27	376
7	HSSGE07	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	84	1285	1	1271	15	15	207	1	28	29	207
8	HMBX59	97897 02/26/97 209043 05/15/97	pBluescript	18	436	87	384	157	157	141	1	21	22	42

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
9	HNGIT22	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	19	503	1	503	23	142	1	19	20	40
10	HERAD57	97897 02/26/97 209043 05/15/97	Uni-ZAP XR	20	358	1	358	147	143	1	31	32	69
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	21	1926	573	1926	157	144	1	30	31	482
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	85	394	1	394	166	208	1	20	21	23
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	86	1925	573	1925	157	209	1	30	31	482
11	HCENJ40	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	87	1818	30	1298	1137	210	1			12

Gene No.	cDNA Clone ID	ATCC Deposit No. Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
12	HCSRA90	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	22	1224	64	557	80	80	145	1	30	31	225
13	HBJFC03	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	23	694	1	694	181	181	146	1	39	40	44
13	HBJFC03	97898 02/26/97 209044 05/15/97	Uni-ZAP XR	88	539	1	539	215	215	211	1	18	19	19
14	HSNBL85	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	24	796	405	796	1	1	147	1	30	31	131
14	HSNBL85	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	89	855	300	855	513	513	212	1	37	38	54
15	HTEY26	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	25	662	205	653	77	77	148	1	30	31	91

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
15	HTEBY26	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	90	628	198	625		275	213	1	31	32	34
16	HMAH107	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	26	1105	40	1105	88	88	149	1	18	19	164
16	HMAH107	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	91	1053	61	1009	79	79	214	1	22	23	229
17	HSKNY94	97899 02/26/97 209045 05/15/97	pBluescript	27	1017	1	1017	97	97	150	1	30	31	138
17	HSKNY94	97899 02/26/97 209045 05/15/97	pBluescript	93	2492	1	943	100	100	216	1	27	28	126
18	HMCDA67	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	28	391	1	391	169	169	151	1	29	30	57

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
19	HOSFF45	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	29	1139	6	1139	109	109	152	1	44	45	47
19	HOSFF45	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	94	3058	1795	2847	1868	1868	217	1	46	47	46
20	HMJAA51	97899 02/26/97 209045 05/15/97	pSport1	30	465	1	370	47	47	153	1	28	29	41
20	HMJAA51	97899 02/26/97 209045 05/15/97	pSport1	95	1099	664	1000	669	669	218	1	33	34	40
21	HTEBF05	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	31	702	1	702	403	403	154	1	24	25	71
22	HTEAL31	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	32	1142	1	518	49	49	155	1	47	48	105

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
22	HTEAI.31	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	96	1580	23	422	32	32	219	1	47	48	104
23	HBMCT32	97899 02/26/97 209045 05/15/97	pBluescript	33	928	1	928	48	48	156	1	27	28	28
23	HBMCT32	97899 02/26/97 209045 05/15/97	pBluescript	97	678	72	593	89	89	220	1	27	28	28
24	HSKXE91	97899 02/26/97 209045 05/15/97	pBluescript	34	773	1	773	39	39	157	1	22	23	52
24	HSKXE91	97899 02/26/97 209045 05/15/97	pBluescript	98	1253	507	1253	507	507	221	1			16
25	HPWTR39	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	35	453	1	453	40	40	158	1	25	26	74

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
26	HTLEV12	97899 02/26/97 209045 05/15/97	Uni-ZAP XR	36	459	1	459	25	25	159	1	24	25	80
27	HSPAF93	97900 02/26/97 209046 05/15/97	pSport1	37	509	1	509	1	1	160	1	19	20	138
27	HSPAF93	97900 02/26/97 209046 05/15/97	pSport1	99	447	1	447	7	7	222	1	23	24	137
28	HHFGL62	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	38	598	1	598	1	1	161	1	21	22	177
28	HHFGL62	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	100	611	37	611	17	17	223	1	26	27	49
29	HCE1U14	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	39	454	1	454	1	1	162	1	21	22	71

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
29	HCE1114	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	101	609	176	609	237	237	224	1			14
30	HEBDA39	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	40	425	1	376	223	223	163	1	18	19	66
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	41	2471	141	2471	213	213	164	1	30	31	154
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	102	1770	47	1721	119	119	225	1	31	32	154
31	HTHBA79	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	103	1832	96	1777	138	138	226	1			9
32	HAGBB70	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	42	2659	1172	2659	119	119	165	1	18	19	103

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
32	HAGBB70	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	104	2237	878	2237	1134	1134	227	1			19
33	HETDG84	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	43	1635	100	1580	299	299	166	1	20	21	80
34	HTEGA81	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	44	780	19	717	10	10	167	1	23	24	92
34	HKGAJ40	209236 09/04/97	pSport1	105	1822	1	1023	272	272	228	1	23	24	93
34	HKMLK44	209084 05/29/97	pBluescript	106	1712	1	1669	168	168	229	1	21	22	93
35	HTXAK60	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	45	2378	1337	2378	1437	1437	168	1	30	31	57
35	HTXAK60	97900 02/26/97 209046 05/15/97	Uni-ZAP XR	107	1969	1068	1892	989	989	230	1	23	24	36

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
36	HMHRN40	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	46	1772	69	1772	129	129	169	1	30	31	231
36	HMHRN40	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	108	1734	65	1734	100	100	231	1	29	30	80
37	HFVGS85	97901 02/26/97 209047 05/15/97	pBluescript	47	1107	70	1107	83	83	170	1	30	31	71
38	HERAH81	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	48	805	167	764	167	167	171	1	23	24	64
39	HMSEU04	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	49	1408	131	1258	364	364	172	1	22	23	74
40	HNEDJ57	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	50	1813	1	1184	2	2	173	1	1	2	333

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
41	HNTME13	97901 02/26/97 209047 05/15/97	pSport1	51	2070	74	2070	142	142	174	1	20	21	195
41	HNTME13	97901 02/26/97 209047 05/15/97	pSport1	109	2003	15	1957	68	68	232	1	22	23	300
42	HSXBI25	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	52	1426	1	1426	158	158	175	1	25	26	264
42	HSXBI25	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	110	1320	80	1311	41	41	233	1	29	30	312
43	HSXCK41	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	53	1720	1	1720	161	161	176	1	22	23	137
43	HSXCK41	97901 02/26/97 209047 05/15/97	Uni-ZAP XR	111	1962	299	1962		566	234	1	33	34	47

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
44	HE8CJ26	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	54	1117	1	1107	218	218	177	1	25	26	178
44	HE8CJ26	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	112	1785	30	1087		225	235	1	23	24	33
45	HTTDS54	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	55	1903	1	1903	119	119	178	1	31	32	154
45	HTTDS54	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	113	1842	1	1832	80	80	236	1	36	37	312
46	HLHDY31	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	56	1860	133	1838	124	124	179	1	24	25	294
46	HLHDY31	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	114	1960	90	1960	165	165	237	1	24	25	295

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
47	HMCRP63	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	57	1259	320	1010	352	352	180	1	26	27	255
48	HEMGE83	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	58	1186	33	557	12	12	181	1	18	19	323
49	HHSDC22	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	59	428	1	304	172	172	182	1	34	35	46
50	HHSDZ57	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	60	501	1	501	40	40	183	1	62	63	92
50	HHSDZ57	97902 02/26/97 209048 05/15/97	Uni-ZAP XR	115	536	73	536	73	73	238	1	22	23	91
52	HMMAB12	97903 02/26/97 209049 05/15/97	pSport1	62	595	1	595	308	308	185	1	29	30	42

Gene No.	cDNA Clone ID	ATCC Deposit No.: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
52	HMMAR12	97903 02/26/97 209049 05/15/97	pSport1	118	453	1	453	198	198	241	1	26	27	27
53	HSKDW02	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	63	1478	40	1436	176	176	186	1	39	40	58
53	HSKDW02	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	119	2016	211	1957	317	317	242	1	25	26	57
54	HETGL41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	64	2033	1	2033	30	30	187	1	30	31	187
54	HETGL41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	120	2136	110	2134	296	296	243	1	23	24	122
55	HODAZ50	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	65	440	1	440	1	1	188	1	26	27	145

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F.
55	HODAZ50	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	121	219	1	219		1	244	1	10	11	72
56	HSDGE59	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	66	3301	349	1478	341	341	189	1	30	31	83
57	HE6ES13	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	67	1535	1	1535	331	331	190	1	26	27	57
57	HE6ES13	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	122	1686	239	1678		367	245	1	27	28	48
58	HISSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	68	1244	402	1244	57	57	191	1	30	31	310
58	HISSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	123	1211	1	1211	80	80	246	1	30	31	338

Gene No.	cDNA Clone ID	ATCC Deposit No; Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
58	HSSEP68	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	124	1804	402	1526	501	501	247	1			17
59	HRDEV41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	69	1292	1	1278	70	70	192	1	28	29	317
59	HRDEV41	97903 02/26/97 209049 05/15/97	Uni-ZAP XR	125	1282	31	1088	70	70	248	1	21	22	338
60	HILCJ01	97903 02/26/97 209049 05/15/97	pBluescript SK-	70	1031	498	1031	536	536	193	1	30	31	52
61	HSATP28	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	71	855	178	855	187	187	194	1	28	29	41
62	HIIIFGL41	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	72	1274	58	1274	118	118	195	1	42	43	101

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
62	HHFGI41	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	126	1296	88	1237	133	133	249	1	39	40	95
63	HRJEM49	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	73	688	1	688	173	173	196	1	18	19	44
63	HRJEM49	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	127	737	1	737	174	174	250	1	20	21	78
64	HSLDJ95	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	74	1890	1	1890	112	112	197	1	21	22	354
64	HSLDJ95	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	128	1925	1	1829	87	87	251	1	23	24	353
65	HSREG44	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	75	1133	408	1133	531	531	198	1	18	19	73

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
66	HTXCT40	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	76	585	1	585	1	1	199	1	69	70	112
66	HTXCT40	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	129	2713	2023	2713	2133	2133	252	1	39	40	108
67	HRGDF73	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	77	577	1	577	51	51	200	1	23	24	122
68	HRDBF52	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	78	2278	1458	1935	25	25	201	1	23	24	314
68	HRDBF52	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	130	1011	479	1011	701	701	253	1	20	21	44
68	HKMND45	97976 04/04/97	pBluescript	131	2278	1	1929	25	25	254	1	27	28	314
69	HPEBD70	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	79	1143	601	1097	95	95	202	1	6	7	235

83

Gene No.	cDNA Clone ID	ATCC Deposit No: Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	Predicted First AA of Secreted Portion	Last AA of OR F
69	HPERD70	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	132	1088	535	1043	588	588	255	1	27	28	52
70	HMCAB89	97904 02/26/97 209050 05/15/97	Uni-ZAP XR	80	557	1	557	132	132	203	1	25	26	92

83

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X and the translated SEQ ID NO:Y are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used to generate antibodies which bind specifically to the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the mature form, or may be a part of a larger protein, such as a fusion protein (see below).

It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

- 5 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources
- 10 using antibodies of the invention raised against the secreted protein in methods which are well known in the art.

Signal Sequences

- Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch,
- 15 Virus Res. 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, *Nucleic Acids Res.* 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1
- 20 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra.*) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

- In the present case, the deduced amino acid sequence of the secreted polypeptide
- 25 was analyzed by a computer program called SignalP (Henrik Nielsen et al., *Protein Engineering* 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results
- 30 shown in Table 1.

- As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., +
- 35 or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely

uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

10 **Polynucleotide and Polypeptide Variants**

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

15 "Identity" per se has an art-recognized meaning and can be calculated using published techniques. (See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, (1988); BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, (1993); COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, (1994); SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, (1987); and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, (1991).) While there exists a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans. (Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).) Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in "Guide to Huge Computers," Martin J. Bishop, ed., Academic Press, San Diego, (1994), and Carillo, H., and Lipton, D., SIAM J Applied Math 48:1073 (1988).

30 Methods for aligning polynucleotides or polypeptides are codified in computer programs, including the GCG program package (Devereux, J., et al., Nucleic Acids Research (1984) 12(1):387 (1984)), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., J. Molec. Biol. 215:403 (1990), Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711 (using the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2:482-489 (1981)).

35

When using any of the sequence alignment programs to determine whether a particular sequence is, for instance, 95% identical to a reference sequence, the parameters are set so that the percentage of identity is calculated over the full length of the reference polynucleotide and that gaps in identity of up to 5% of the total number of nucleotides in the reference polynucleotide are allowed.

A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245 (1990).) The term "sequence" includes nucleotide and amino acid sequences. In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB search of a DNA sequence to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, and Cutoff Score=1. Gap Penalty=5, Gap Size Penalty 0.05, and Window Size=500 or query sequence length in nucleotide bases, whichever is shorter. Preferred parameters employed to calculate percent identity and similarity of an amino acid alignment are: Matrix=PAM 150, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty=0.05, and Window Size=500 or query sequence length in amino acid residues, whichever is shorter.

As an illustration, a polynucleotide having a nucleotide sequence of at least 95% "identity" to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone, means that the polynucleotide is identical to a sequence contained in SEQ ID NO:X or the cDNA except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the total length (not just within a given 100 nucleotide stretch). In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to SEQ ID NO:X or the deposited clone, up to 5% of the nucleotides in the sequence contained in SEQ ID NO:X or the cDNA can be deleted, inserted, or substituted with other nucleotides. These changes may occur anywhere throughout the polynucleotide.

Further embodiments of the present invention include polynucleotides having at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to a sequence contained in SEQ ID NO:X or the cDNA contained in the deposited clone. Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the polynucleotides having at least 85%, 90%, 95%, 96%, 97%, 98%, or 99% identity

will encode a polypeptide identical to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference polypeptide, is intended that the amino acid sequence of the polypeptide is identical to the reference polypeptide except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the total length of the reference polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

Further embodiments of the present invention include polypeptides having at least 80% identity, more preferably at least 85% identity, more preferably at least 90% identity, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone. Preferably, the above polypeptides should exhibit at least one biological activity of the protein.

In a preferred embodiment, polypeptides of the present invention include polypeptides having at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98%, or 99% similarity to an amino acid sequence contained in SEQ ID NO:Y or the expressed protein produced by the deposited clone.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an

organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level.

Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

5 Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988
10 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

 Moreover, ample evidence demonstrates that variants often retain a biological
15 activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible
20 amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

25 Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form
30 are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

 Thus, the invention further includes polypeptide variants which show
35 substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make

phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

5 The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid
10 substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham
15 and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the
20 protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues
25 Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp. and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues,
30 where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino
35 acids, such as an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Polynucleotide and Polypeptide Fragments

10 In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence contained in the deposited clone or shown in SEQ ID NO:X. The short nucleotide fragments are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length. A fragment "at least 20 nt in
15 length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in the deposited clone or the nucleotide sequence shown in SEQ ID NO:X. These nucleotide fragments are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

20 Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments having a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, and 701 to the end of SEQ ID NO:X or the cDNA contained in the deposited clone. In this context "about" includes the particularly
25 recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity.

In the present invention, a "polypeptide fragment" refers to a short amino acid sequence contained in SEQ ID NO:Y or encoded by the cDNA contained in the
30 deposited clone. Protein fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, and 161 to the end of the coding
35 region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about"

includes the particularly recited ranges, larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotide fragments encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotide fragments encoding these domains are also contemplated.

Other preferred fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Epitopes & Antibodies

In the present invention, "epitopes" refer to polypeptide fragments having antigenic or immunogenic activity in an animal, especially in a human. A preferred embodiment of the present invention relates to a polypeptide fragment comprising an epitope, as well as the polynucleotide encoding this fragment. A region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." In contrast, an "immunogenic epitope" is defined as a part of a protein that elicits an antibody response. (See, for instance, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).)

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, R. A., Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985) further described in U.S. Patent No. 4,631,211.)

5 In the present invention, antigenic epitopes preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between about 15 to about 30 amino acids. Antigenic epitopes are useful to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe, J. G. et al., Science 219:660-666 (1983).)

10 Similarly, immunogenic epitopes can be used to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow, M. et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle, F. J. et al., J. Gen. Virol. 66:2347-2354 (1985).) A preferred immunogenic epitope includes the secreted protein. The immunogenic epitopes may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if
15 it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting.)

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is
20 meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred,
25 as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

Fusion Proteins

30 Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular
35 locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

5 Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the
10 polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

 Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of
15 immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4- polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827: Traunecker et al., Nature 331:84-86
20 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

 Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion
25 proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,
30 would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol.
35 Chem. 270:9459-9471 (1995).)

 Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In

preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the claimed invention.

Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS,

293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and p

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes
5 known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention
10 can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic
15 cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can
20 be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using
25 fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to
30 mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross
35 hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage

analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Both methods rely on binding of the polynucleotide to DNA or RNA. For these techniques, preferred polynucleotides are usually 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the

present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute
5 biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags"
10 which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an
15 individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely
20 small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from
25 polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the
30 present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present
35 invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I , ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I , ^{112}In , $^{99\text{m}}\text{Tc}$), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20

millicuries of ^{99m}Tc . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments," (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder.

Moreover, polypeptides of the present invention can be used to treat disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B), to inhibit the activity of a polypeptide (e.g., an oncogene), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention could be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Biological Activities

The polynucleotides and polypeptides of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules

may be involved in the diseases associated with the biological activity. Thus, the polynucleotides and polypeptides could be used to treat the associated disease.

Immune Activity

5 A polypeptide or polynucleotide of the present invention may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells
10 from pluripotent stem cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotide or polypeptide of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

15 A polynucleotide or polypeptide of the present invention may be useful in treating or detecting deficiencies or disorders of hematopoietic cells. A polypeptide or polynucleotide of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat those disorders associated with a decrease in certain (or many) types hematopoietic
20 cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein disorders (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

25 Moreover, a polypeptide or polynucleotide of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotide or polypeptide of the present invention could be used to treat blood
30 coagulation disorders (e.g., afibrinogenemia, factor deficiencies), blood platelet disorders (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotide or polypeptide of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment of heart attacks
35 (infarction), strokes, or scarring.

A polynucleotide or polypeptide of the present invention may also be useful in treating or detecting autoimmune disorders. Many autoimmune disorders result from

inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation,
5 differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune disorders.

Examples of autoimmune disorders that can be treated or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis,
10 glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitus, and autoimmune
15 inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated by a polypeptide or polynucleotide of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

20 A polynucleotide or polypeptide of the present invention may also be used to treat and/or prevent organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The
25 administration of a polypeptide or polynucleotide of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polypeptide or polynucleotide of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide may
30 inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat inflammatory conditions, both chronic and acute conditions, including inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute
35 rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

A polypeptide or polynucleotide can be used to treat or detect hyperproliferative disorders, including neoplasms. A polypeptide or polynucleotide of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polypeptide or polynucleotide of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative disorders can be treated. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating hyperproliferative disorders, such as a chemotherapeutic agent.

Examples of hyperproliferative disorders that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but are not limited to neoplasms located in the: abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative disorders can also be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of such hyperproliferative disorders include, but are not limited to: hypergammaglobulinemia, lymphoproliferative disorders, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenström's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

Infectious Disease

A polypeptide or polynucleotide of the present invention can be used to treat or detect infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, the polypeptide or polynucleotide of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

- Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention. Examples of viruses, include, but are not limited to the following DNA and RNA viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza), Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiolitis, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. A polypeptide or polynucleotide of the present invention can be used to treat or detect any of these symptoms or diseases.
- Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following Gram-Negative and Gram-positive bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia, Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, Enterobacteriaceae (Klebsiella, Salmonella, Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Pasteurellaceae Infections (e.g., Actinobacillus, Haemophilus, Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, and Staphylococcal. These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis, Chlamydia, Syphilis, Diphtheria.

Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. A polypeptide or polynucleotide of the present invention can be used to treat or detect
5 any of these symptoms or diseases.

Moreover, parasitic agents causing disease or symptoms that can be treated or detected by a polynucleotide or polypeptide of the present invention include, but not limited to, the following families: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis,
10 Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas. These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), Malaria, pregnancy complications, and toxoplasmosis. A polypeptide or polynucleotide
15 of the present invention can be used to treat or detect any of these symptoms or diseases.

Preferably, treatment using a polypeptide or polynucleotide of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide
20 of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

25 A polynucleotide or polypeptide of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal
30 disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vascular (including vascular endothelium), nervous, hematopoietic, and
35 skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic disorders (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated using the polynucleotide or polypeptide of the present invention.

Chemotaxis

A polynucleotide or polypeptide of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide of the present invention may increase chemotaxis activity of particular cells. These chemotactic molecules can then be used to treat inflammation, infection, hyperproliferative disorders, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxis molecules can be used to treat wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotaxis molecules of the present invention can also attract fibroblasts, which can be used to treat wounds.

It is also contemplated that a polynucleotide or polypeptide of the present invention may inhibit chemotaxis activity. These molecules could also be used to treat

disorders. Thus, a polynucleotide or polypeptide of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

5 A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or
10 small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural
15 receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell
20 membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

25 The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product
30 mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

35 Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The

antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptide from suitably manipulated cells or tissues.

Therefore, the invention includes a method of identifying compounds which bind to a polypeptide of the invention comprising the steps of: (a) incubating a candidate binding compound with a polypeptide of the invention; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with a polypeptide of the invention, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Other Activities

A polypeptide or polynucleotide of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

A polypeptide or polynucleotide of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, a polypeptide or polynucleotide of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

A polypeptide or polynucleotide of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive disorders), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

A polypeptide or polynucleotide of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining

whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid
5 molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

10 A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any
15 integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide
20 sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein
25 identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide
30 sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least
35 two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the

amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at
5 least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at
10 least 95% identical to the amino acid sequence of the secreted portion of the protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a
15 polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in
20 the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1;
25 and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining
30 whether the sequence of said polypeptide molecule in said sample is at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of
35 polypeptides in said sample to an antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an

amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- 5 Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

- Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in
10 said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded
15 by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

- Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a
20 sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

- Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample
25 obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid
30 sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

 In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

- 35 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least

90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated

polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited Plasmid</u>
	Lambda Zap	pBluescript (pBS)
20	Uni-Zap XR	pBluescript (pBS)
	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSPORT 2.0	pCMVSPORT 2.0
25	pCMVSPORT 3.0	pCMVSPORT 3.0
	pCR [®] 2.1	pCR [®] 2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Altting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Altting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS-. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which

are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSPORT 2.0 and pCMVSPORT 3.0, were obtained from
5 Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lacmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1
10 Blue. Vector pCR[®]2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention does not comprise the
15 phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing
20 the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

25 Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized
30 using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with ³²P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as
35 XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above.

The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 μ l of reaction mixture with 0.5 μ g of the above cDNA template. A convenient reaction mixture is 1.5-5 mM $MgCl_2$, 0.01% (w/v) gelatin, 20 μ M each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

Several methods are available for the identification of the 5' or 3' non-coding portions of a gene which may not be present in the deposited clone. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3' "RACE" protocols which are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then

be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

- 5 This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that
- 10 the 5' end sequence belongs to the desired gene.

Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide

- 15 A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

Example 3: Tissue Distribution of Polypeptide

- 20 Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with P³² using the rediprime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.),
- 25 according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

- Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's
- 30 protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and the films developed according to standard procedures.

Example 4: Chromosomal Mapping of the Polynucleotides

- 35 An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This

primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95°C: 1 minute, 56°C: 1 minute, 70°C. This cycle is repeated 32 times followed by one 5 minute cycle at 70°C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios, Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

10 **Example 5: Bacterial Expression of a Polypeptide**

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4°C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8. the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4°C or frozen at -80°C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number XXXXXX.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an E. coli origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or

Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

5

Example 6: Purification of a Polypeptide from an Inclusion Body

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

- 10 Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

- 15 The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

- 20 The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

- 25 Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

30 To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive

Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

5 Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium
10 acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

15 The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commaissie blue stained 16% SDS-PAGE gel when 5 μ g of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

20

Example 7: Cloning and Expression of a Polypeptide in a Baculovirus Expression System

In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong
25 polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the
30 same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

35 Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription,

translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

Specifically, the cDNA sequence contained in the deposited clone, including the
5 AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al.,
10 "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

15 The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4
20 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA
25 sequencing.

Five μ g of a plasmid containing the polynucleotide is co-transfected with 1.0 μ g of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by
Felgner et al., *Proc. Natl. Acad. Sci. USA* 84:7413-7417 (1987). One μ g of
30 BaculoGold™ virus DNA and 5 μ g of the plasmid are mixed in a sterile well of a microtiter plate containing 50 μ l of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μ l Lipofectin plus 90 μ l Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm
35 tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27°C. The transfection solution is then removed from the plate

and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ³⁵S-methionine and 5 µCi ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 8: Expression of a Polypeptide in Mammalian Cells

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV, HIV and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used
5 include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable
10 marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of
15 interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991). Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the
20 mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the
25 expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession No. 209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., Cell 41:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the
30 cloning of the gene of interest. The vectors also contain the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene, and the mouse DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate restriction enzymes and then dephosphorylated using calf intestinal phosphates by
35 procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the protocol outlined in Example 1. If the naturally occurring signal sequence is used to produce the

secreted protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

5 The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then
10 transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transfection. Five μ g of the expression plasmid pC6 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo
15 contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418.
20 After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same
25 procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 9: Protein Fusions

30 The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example 5; see also EP A 394.827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and
35 albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the

activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which
5 outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an
10 expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated
15 by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally
20 occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

Human IgG Fc region:

```
GGGATCCGGAGCCCAAATCTTCTGACAAACTCACACATGCCCACCGTGCC  
25 CAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAC  
CAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGGTGGT  
GGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACG  
GCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAAC  
AGCACGTACCGTGTGGTCAGCGTCCTACCGTCCTGCACCAGGACTGGCTG  
30 AATGGCAAGGAGTACAAGTGAAGGTCTCCAACAAAGCCCTCCCAACCCCC  
ATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGT  
GTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCT  
GACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTGGAGTGGGA  
GAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGG  
35 ACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCA  
GGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGC
```

ACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATGAGTGC
GACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

Example 10: Production of an Antibody from a Polypeptide

5 The antibodies of the present invention can be prepared by a variety of methods.
(See, Current Protocols, Chapter 2.) For example, cells expressing a polypeptide of
the present invention is administered to an animal to induce the production of sera
containing polyclonal antibodies. In a preferred method, a preparation of the secreted
protein is prepared and purified to render it substantially free of natural contaminants.
10 Such a preparation is then introduced into an animal in order to produce polyclonal
antisera of greater specific activity.

 In the most preferred method, the antibodies of the present invention are
monoclonal antibodies (or protein binding fragments thereof). Such monoclonal
antibodies can be prepared using hybridoma technology. (Köhler et al., Nature
15 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J.
Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell
Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures
involve immunizing an animal (preferably a mouse) with polypeptide or, more
preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in
20 any suitable tissue culture medium; however, it is preferable to culture cells in Earle's
modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at
about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about
1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

 The splenocytes of such mice are extracted and fused with a suitable myeloma
25 cell line. Any suitable myeloma cell line may be employed in accordance with the
present invention; however, it is preferable to employ the parent myeloma cell line
(SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are
selectively maintained in HAT medium, and then cloned by limiting dilution as
described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells
30 obtained through such a selection are then assayed to identify clones which secrete
antibodies capable of binding the polypeptide.

 Alternatively, additional antibodies capable of binding to the polypeptide can be
produced in a two-step procedure using anti-idiotypic antibodies. Such a method
makes use of the fact that antibodies are themselves antigens, and therefore, it is
35 possible to obtain an antibody which binds to a second antibody. In accordance with
this method, protein specific antibodies are used to immunize an animal, preferably a

mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

Example 11: Production Of Secreted Protein For High-Throughput Screening Assays

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at 2×10^5 cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

5 The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a
10 multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

15 Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel,
20 adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37°C for 6 hours.

While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (see below) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock
25 solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37°C for 45 or 72 hours
30 depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays
35 described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing expression of other

proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

5 HGS-CHO-5 medium formulation:

Inorganic Salts

CaCl ₂ (anhyd)	116.6 mg/L
CuSO ₄ ·5H ₂ O	0.00130
Fe(NO ₃) ₃ ·9H ₂ O	0.050
FeSO ₄ ·7H ₂ O	0.417
KCl	311.80
MgCl ₂	28.64
MgSO ₄	48.84
NaCl	6995.50
NaHCO ₃	2400.0
NaH ₂ PO ₄ ·H ₂ O	62.50
Na ₂ HPO ₄	71.02
ZnSO ₄ ·7H ₂ O	.4320

Lipids

Arachidonic Acid	.002 mg/L
Cholesterol	1.022
DL-alpha-Tocopherol-Acetate	.070
Linoleic Acid	0.0520
Linolenic Acid	0.010
Myristic Acid	0.010
Oleic Acid	0.010
Palmitric Acid	0.010
Palmitic Acid	0.010
Pluronic F-68	100
Stearic Acid	0.010
Tween 80	2.20

10 Carbon Source

D-Glucose	4551 mg/L
-----------	-----------

Amino Acids

L- Alanine	130.85 mg/ml
L-Arginine-HCL	147.50
L-Asparagine-H ₂ O	7.50

L-Aspartic Acid	6.65
L-Cystine-2HCL-H ₂ O	29.56
L-Cystine-2HCL	31.29
L-Glutamic Acid	7.35
L-Glutamine	365.0
Glycine	18.75
L-Histidine-HCL-H ₂ O	52.48
L-Isoleucine	106.97
L-Leucine	111.45
L-Lysine HCL	163.75
L-Methionine	32.34
L-Phenylalanine	68.48
L-Proline	40.0
L-Serine	26.25
L-Threonine	101.05
L-Tryptophan	19.22
L-Tyrosine-2Na-2H ₂ O	91.79
L-Valine	99.65

Vitamins

Biotin	0.0035 mg/L
D-Ca Pantothenate	3.24
Choline Chloride	11.78
Folic Acid	4.65
i-Inositol	15.60
Niacinamide	3.02
Pyridoxal HCL	3.00
Pyridoxine HCL	0.031
Riboflavin	0.319
Thiamine HCL	3.17
Thymidine	0.365
Vitamin B ₁₂	0.680

Other Components

HEPES Buffer	25 mM
Na Hypoxanthine	2.39 mg/L
Lipoic Acid	0.105
Sodium Putrescine-2HCL	0.081
Sodium Pyruvate	55.0
Sodium Selenite	0.0067
Ethanolamine	20uM
Ferric Citrate	0.122
Methyl-B-Cyclodextrin complexed with Linoleic Acid	41.70

Methyl-B-Cyclodextrin complexed with Oleic Acid	33.33
Methyl-B-Cyclodextrin complexed with Retinal Acetate	10

Adjust osmolarity to 327 mOsm

Example 12: Construction of GAS Reporter Construct

5 One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the Jaks-STATs pathway bind to gamma activation site "GAS" elements or interferon-sensitive responsive element ("ISRE"), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

10 GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or "STATs." There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I. cells after treatment with
15 IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase ("Jaks")
20 family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51
25 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a
30 WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

- 5 Therefore, activation of the Jaks-STATs pathway, reflected by the binding of the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

	Ligand	tyk2	JAKs Jak1	Jak2	Jak3	STATs	GAS(elements) or ISRF
	<u>IFN family</u>						
5	IFN- α /B	+	+	-	-	1,2,3	ISRE
	IFN- γ		+	+	-	1	GAS (IRF1>Lys6>IFP)
	IL-10	+	?	?	-	1,3	
	<u>gp130 family</u>						
10	IL-6 (Pleiotrohic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
	IL-11(Pleiotrohic)	?	+	?	?	1,3	
	OnM(Pleiotrohic)	?	+	+	?	1,3	
	LIF(Pleiotrohic)	?	+	+	?	1,3	
	CNTF(Pleiotrohic)	-/+	+	+	?	1,3	
15	G-CSF(Pleiotrohic)	?	+	?	?	1,3	
	IL-12(Pleiotrohic)	+	-	+	+	1,3	
	<u>g-C family</u>						
20	IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
	IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
	IL-7 (lymphocytes)	-	+	-	+	5	GAS
	IL-9 (lymphocytes)	-	+	-	+	5	GAS
	IL-13 (lymphocyte)	-	+	?	?	6	GAS
	IL-15	?	+	?	+	5	GAS
25	<u>gp140 family</u>						
	IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
	IL-5 (myeloid)	-	-	+	-	5	GAS
	GM-CSF (myeloid)	-	-	+	-	5	GAS
30	<u>Growth hormone family</u>						
	GH	?	-	+	-	5	
	PRL	?	+/-	+	-	1,3,5	
	EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
35	<u>Receptor Tyrosine Kinases</u>						
	EGF	?	+	+	-	1,3	GAS (IRF1)
	PDGF	?	-	+	-	1,3	
	CSF-1	?	+	+	-	1,3	GAS (not IRF1)
40							

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the IRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

5':GCGCCTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCG
10 AAATGATTTCCTCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTGGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

5':CTCGAGATTTCCTCCGAAATCTAGATTTCCTCCGAAATGATTTCCTCCGAAATG
20 ATTTTCCTCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCC
CTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGC
CCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGC
CTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTT
TGCAAAAAGCTT:3' (SEQ ID NO:5)

25 With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase,
30 alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter
35 element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using Sall and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning
5 site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules
10 containing NFK-B and EGR promoter sequences are described in Examples 15 and 16. However, many other promoters can be substituted using the protocols described in these Examples. For instance, SRE, IL-2, NFAT1, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter
15 construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

Example 13: High-Throughput Screening Assay for T-cell Activity.

The following protocol is used to assess T-cell activity by identifying factors,
20 such as growth factors and cytokines, that may proliferate or differentiate T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and
25 Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately
30 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to
35 generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in RPMI + 10% serum with 1% Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life Technologies)

with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells (10^7 per transfection), and resuspend in OPTI-MEM to a final
5 concentration of 10^7 cells/ml. Then add 1ml of 1×10^7 cells in OPTI-MEM to T25 flask and incubate at 37°C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Gentamicin, and 1% Pen-Strep. These cells are treated with supernatants containing a polypeptide as produced by the protocol described in Example 11.

10 On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

15 Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100,000 cells per well).

After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12
20 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul samples
25 from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophane covers) and stored at -20°C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4°C and serve as a source of material for repeating the assay on a specific well if desired.

30 As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

Example 14: High-Throughput Screening Assay Identifying Myeloid Activity

The following protocol is used to assess myeloid activity by identifying factors, such as growth factors and cytokines, that may proliferate or differentiate myeloid cells.

- 5 Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

- 10 To transiently transfect U937 cells with the GAS/SEAP/Neo construct produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest 2×10^7 U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

- 15 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM KCl, 375 uM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 1 mM MgCl_2 , and 675 uM CaCl_2 . Incubate at 37°C for 45 min.

- 20 Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37°C for 36 hr.

The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400 ug/ml G418 for couple of passages.

- 25 These cells are tested by harvesting 1×10^5 cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of 5×10^5 cells/ml. Plate 200 ul cells per well in the 96-well plate (or 1×10^5 cells/well).

- 30 Add 50 ul of the supernatant prepared by the protocol described in Example 11. Incubate at 37°C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.

When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor). The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., *Oncogene* 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

- 5 The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as 5×10^5 cells/ml.

- 10 Add 200 μ l of the cell suspension to each well of 96-well plate (equivalent to 1×10^5 cells/well). Add 50 μ l supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ μ l of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

15

Example 16: High-Throughput Screening Assay for T-cell Activity

- 20 NF- κ B (Nuclear Factor κ B) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF- κ B regulates the expression of genes involved in immune cell activation, control of apoptosis (NF- κ B appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

- 25 In non-stimulated conditions, NF- κ B is retained in the cytoplasm with I- κ B (Inhibitor κ B). However, upon stimulation, I- κ B is phosphorylated and degraded, causing NF- κ B to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- κ B include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

- 30 Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF- κ B promoter element are used to screen the supernatants produced in Example 11. Activators or inhibitors of NF- κ B would be useful in treating

diseases. For example, inhibitors of NF- κ B could be used to treat those diseases related to the acute or chronic activation of NF- κ B, such as rheumatoid arthritis.

To construct a vector containing the NF- κ B promoter element, a PCR based strategy is employed. The upstream primer contains four tandem copies of the NF- κ B binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCGGGGACTTTCCGGGACTTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5':CTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGGACTTTCC
ATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCCA
TCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACT
AATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTC
CAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:
3' (SEQ ID NO:10)

Next, replace the SV40 minimal promoter element present in the pSEAP2-promoter plasmid (Clontech) with this NF- κ B/SV40 fragment using XhoI and HindIII. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

In order to generate stable mammalian cell lines, the NF- κ B/SV40/SEAP cassette is removed from the above NF- κ B/SEAP vector using restriction enzymes Sall and NotI, and inserted into a vector containing neomycin resistance. Particularly, the NF- κ B/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the GFP gene, after restricting pGFP-1 with Sall and NotI.

Once NF- κ B/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are created and maintained according to the protocol described in Example 13. Similarly, the method for assaying supernatants with these stable Jurkat T-cells is also described in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to wells H9, H10, and H11, with a 5-10 fold activation typically observed.

Example 17: Assay for SEAP Activity

As a reporter molecule for the assays described in Examples 13-16, SEAP activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 μ l of 2.5x dilution buffer into Optiplates containing 35 μ l of a supernatant. Seal the plates with a plastic sealer and incubate at 65°C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 μ l Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 μ l Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

Reaction Buffer Formulation:

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6

23	125	6.25
24	130	6.5
25	135	6.75
26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-3, used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star black 96-well plate with clear bottom. The plate is incubated in a CO₂ incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-3 is made in 10% pluronic acid DMSO. To load the cells with fluo-3, 50 μ l of 12 μ g/ml fluo-3 is added to each well. The plate is incubated at 37°C in a CO₂ incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 μ l of buffer.

- 5 For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to 2-5x10⁶ cells/ml with HBSS in a 50-ml conical tube. 4 μ l of 1 mg/ml fluo-3 solution in 10% pluronic acid DMSO is added to each ml of cell suspension. The tube is then placed in a 37°C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to 1x10⁶ cells/ml, and dispensed into a microplate, 100
10 μ l/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 μ l, followed by an aspiration step to 100 μ l final volume.

For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-3. The supernatant is added to the well, and a change in fluorescence is detected.

- To measure the fluorescence of intracellular calcium, the FLIPR is set for the
15 following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 μ l. Increased emission at 530 nm indicates an extracellular signaling event which has resulted in an increase in the intracellular Ca⁺⁺ concentration.

20

Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase Activity

- The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase
25 RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

- 30 Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members
35 of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 µl of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4°C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200µl/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 µl of the supernatant produced in Example 11, the medium was removed and 100 µl of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na₃VO₄, 2 mM Na₄P₂O₇ and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4°C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 µm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4°C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a

biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

- 5 The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg₂₊ (5mM ATP/50mM MgCl₂), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride, pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl₂, 5 mM MnCl₂, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the
- 10 components gently and preincubate the reaction mix at 30°C for 2 min. Initiate the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of 120mM EDTA and place the reactions on ice.

- Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction
- 15 mixture to a microtiter plate (MTP) module and incubating at 37°C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times. Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37°C for one hour. Wash the well as
- 20 above.

- Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of
- 25 tyrosine kinase activity.

Example 20: High-Throughput Screening Assay Identifying Phosphorylation Activity

- As a potential alternative and/or complement to the assay of protein tyrosine
- 30 kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase,
- 35 Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other

phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1 μ g/ml) for 2 hr at room temp. (RT). The plates are then
5 rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules, this step can easily be modified by substituting a monoclonal antibody detecting any of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4°C
10 until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal medium (DMEM) and then treated with EGF (6ng/well) or 50 μ l of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts
15 filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1 μ g/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and
20 Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

25

Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from
30 these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C; and 60-120 seconds at 70°C, using buffer solutions described in Sidransky, D., et al., Science 252:706 (1991).

35 PCR products is then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies).

The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton, T.A. and Graham, M.W., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals is identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenin deoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson, Cg. et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson, Cg. et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

A polypeptide of the present invention can be detected in a biological sample. and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect soluble polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method

described in Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 23: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracisternally, intravaginally,

intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15:167-277 (1981), and R. Langer, *Chem. Tech.* 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al.) or poly-D-(-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., *Proc. Natl. Acad. Sci. USA* 82:3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. USA* 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Example 24: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form.

- 5 Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily
10 dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided in Example 23.

Example 25: Method of Treating Increased Levels of the Polypeptide

- 15 Antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5,
20 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

Example 26: Method of Treatment Using Gene Therapy

- 25 One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is
30 turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin, is added. The flasks are then incubated at 37°C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

5 pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set
10 forth in Example 1. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to
15 transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is
20 then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media,
25 containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is
30 required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is being produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

- 5 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference.

(1) GENERAL INFORMATION:

(i) APPLICANTS: Human Genome Sciences, Inc. et al.

(ii) TITLE OF INVENTION: 70 Human Secreted Proteins

5 (iii) NUMBER OF SEQUENCES: 273

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Human Genome Sciences, Inc.

(B) STREET: 9410 Key West Avenue

(C) CITY: Rockville

10 (D) STATE: Maryland

(E) COUNTRY: USA

(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

15 (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

(B) COMPUTER: HP Vectra 486/33

(C) OPERATING SYSTEM: MSDOS version 6.2

(D) SOFTWARE: ASCII Text

(vi) CURRENT APPLICATION DATA:

20 (A) APPLICATION NUMBER:

(B) FILING DATE: March 6, 1998

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

25 (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: A. Anders Brookes

(B) REGISTRATION NUMBER: 36,373

(C) REFERENCE/DOCKET NUMBER: PS001PCT

160

(vi) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (301) 309-8504

(E) TELEFAX: (301) 309-8439

5 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 733 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGGATCCGGA GCCCAAATCT TCTGACAAA CTCACACATG CCCACCGTGC CCAGCACCTG 60
 AATTGGAGGG TGGACCGTCA GTCTTCCTCT TCCCCC AAA ACCCAAGGAC ACCCTCATGA 120
 15 TCTCCCGGAC TCTTGAAGTC ACATGCGTGG TGCTGGACGT AAGCCACGAA GACCCGTAGG 180
 TCAAGTTCAA CTGTAAGTG GACGGCGTGG AGGTGCATAA TGGCAAGACA AAGCCGCGGG 240
 AGGAGCAGTA CAACAGCAGG TACCGTGTGG TCAGCGTCTT CACCGTCTTG CACCAGGACT 300
 GGCTGAATGG CAAGGAGTAC AAGTGCAAGG TCTCCAACAA AGCCCTCCCA ACCCCCATCG 360
 AGAAAACCAT CTGCAAGGC AAAGGGCAGC CCGGAGAACC ACACTGTATC ACCCTGCCCC 420
 20 CATCCCGGGA TGAGTTGACC AAGAACCAGG TCAGCCTGAC CTGCTTGGTC AAAGGCTTCT 480
 ATCCAAGCGA CATGCGCTG GACTGGGAGA GGAATGGGCA GCGGAGAAC AACTACAAGA 540
 CCACGCTCTC CGTGTGTGAC TCCGACGGCT CTTCTTCTCT CTACAGCAAG CTCACCGTGG 600
 ACAAGAGCAG CTGTAAGCAG GGAACGTCT TCTCATGCTC GTGATGCAT GAGGCTCTGC 660
 ACAACCACTA CACGAGAGAG AGCCTCTCCC TGTCTCGGG TAAATGAGTG CGACGGCGGC 720
 25 GACTCTAGAG GAT 733

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

30 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Trp Ser Xaa Trp Ser

5 1 5

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 86 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

15 GCGCCTCGAG ATTTCCTCCGA AATCTAGATT TCCCCGAAAT GATTTCCTCCG AAATGATTTC 60
 CCGGAAATAT CTGCCATCTC AATTAG 86

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

25 GCGGCAAGCT TTTTGCAAAG CCTAGGC 27

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 271 base pairs
 (B) TYPE: nucleic acid

162

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTCGAGATTT CCCCGAAATC TAGATTTCCC CGAAATGATT TCCCCGAAAT GATTCCCCCG 60
5 AAATATCTGC CATCTCAATT AGTCAGCAAC CATACTCCCG CCCCTAACTC CGCCCATCCC 120
GCCCCTAACT CGGCCAGTT CGGCCATTC TCCGCCCAT GGCTGACTAA TTTTTTTTAT 180
TTATGCAGAG GCGAGGCCG CCGGGCCTC TGAGCTATTC CAGAAGTAGT GAGGAGGCTT 240
TTTTGGAGGC CTAGGCTTTT GCAAAAAGCT T 273

10 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCGCTCGAGG GATGACAGCG ATAGAACCCC GG 32

20 (2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCGAAGCTTC GCGACTCCCC GGATCCGCCT C 31

30 (2) INFORMATION FOR SEQ ID NO: 8:

163

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

5

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGGACTTTC CC

12

10 (2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GCGGCCTCGA GGGGACTTTC CCGGGGACTT TCCGGGGACT TTCCGGGACT TTCCATCCTG

60

CCATCTCAAT TAG

73

20

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 256 base pairs

25

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

30 CTCGAGGGGA CTCTCCCGGG GACTTTCGGG GGACTTTCCG GGACTTTCCA TCTGCCATCT

60

CAATTASTCA GCAACCATAG TCCCGCCCT AACTCCGCC ATCCCCCCC TAACTCGGC 120
 CAGTCCGCC CATTCTCCGC CCATGGGTG ACTAATTTT TTATTTATG CAGAGGCCGA 180
 GGGCGGCTCG GCTCTGAGC TACTCCAGAA GTAGTGAGGA GGTCTTTTG GAGGCTAGG 240
 CTTTTCGAAA AAGCTT 256

5

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1739 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGCTCCCGA GGCCGCGGA CCTGCAGAGA GGACAGCCGG CCTCCGCCG GACATGCCGG 60
 15 CCCAGAGCT CCCAGGCTC GGGTCCCGT TGCTGCTGTT GCTCTTCTG CTGCTGCCG 120
 CGCCGCTGT CCTGCCCCA AGGCGCAGC GTTTCGACCC CACCTGGAG TCCCTGGAG 180
 CCGCCAGCT GCGCGCTGG TTGACCCAG CCAAGTTCG CATCTTCAT CACTGGGGAG 240
 TGTTTCCGT GCGCCTTC GGTAGCGAGT GGTCTGGTG GTATTGCAA AAGGAAAAGA 300
 TACCGAAGTA TGTGGAATTT ATGAAAGATA ATTACCTCC TATTTTCAA TATGAAGATT 360
 20 TTGACCACT ATTTACAGCA AAATTTTTA ATGCCAACCA TGGGGCAAT ATTTTTCAG 420
 CCTCTGTGC CAAATACATT GTCTTAACTT CCAACATCA TGAAGGCTT ACCTTGTGG 480
 GGTCAAGATA TTGCTGGAAC TGAATGACA TAGATGAGGG GCGAAGAGG GACATTGTCA 540
 AGGAAGTGA GTTAGCCATT AGGAACAGAA CTGACCTCG TTTGAGCTG TACTATTCCC 600
 TTTTGAATG GTTTCATCCG GTTTCCTTG AGGATGAATC CAGTTCATT CATAAGCGGC 660
 25 AATTTCCAGT TTCTAAGACA TTGCGAGAGT TCTATGAGTT AGTGAACAAT TATCAGCCTG 720
 AGGTTCTGT GTGGAATGT TATGAGAGG CACCGATCA ATACTGAAAT AACACAGGCT 780
 TCTTGGCTG GTTATATAAT GAAAGGCCAG TTGCGGGCAC AGTAGTCAG AATGATCGTT 840
 GGGGAGCTG TAGCATCTGT AAGCATGCTG GCTTCTATAC CTGCAGTGAT CGTTATAACC 900
 CAGGACATCT TTGCCACAT AAATGGGAAA ACTGCATGAC AATAGACAAA CTGTCCTGGG 960
 30 GCTATAGSAG GGAAGCTGGA ATCTGAGCT ATCTTACAAT TGAAGAATTG GTGAAGCAAC 1020

165

TTGTAGAGAC AGTTTCATGT GGAGGAAATC TTTTGATGAA TATTGGGCCC AACTAGATG 1080
 GCACCATTTT TGTAGTTTTT GAGGAGCGAC TGAGGCAAAT GGGGTCTGG CTAAAAGTCA 1140
 ATGGAGAAGC TATTTATGAA ACCCATACCT GCGGATCCCA GAATGACACT GTCACCCAG 1200
 ATGTGTGGTA CACATCCAAG CCTAAAGAAA AATTAGTCTA TGCCATTTTT CTAAATGGC 1260
 5 CCACATCAGG ACAGCTGTTC CTTGGCCATC CCAAAGCTAT TGTGGGGCA ACAGAGGTGA 1320
 AACTACTGGG CCATGGACAG CCACTTAACCT GGATTTCTTT GGAGCAAAAT GGCATTATGG 1380
 TAGAACTGCC ACAGCTAACC ATTATCAGA TGCCGTGTAA ATGGGCTGG GCTCTAGCCC 1440
 TRACTAATGT GATCTAAAGT GCAGCAGAGT GGCTGATGCT GCAAGTTATG TCTAAGGCTA 1500
 GGAACTATCA GGTGTCTATA ATTGTAGCAC ATGGAGAAAG CAAATGTAAA ACTGGATAAG 1560
 10 AAAATTATTT TGGCAGTTCA GCCCTTCCC TTTTCCCAC TAAATTTTTT CTAAATTAC 1620
 CCATGTAACC ATTTTAACTC TCCAGTGCAC TTTGCCATTA AAGTCTCTC ACATTGAAAA 1680
 AAAAAAAAAA AAAAACCCCG GGGGGGGGGC CCGGNACCC CATTTGCCC NTAAAGGGG 1739

(2) INFORMATION FOR SEQ ID NO: 12:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 844 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGCCCCGGG CCCGAGGGG TGGAGCCGG CCGGGCGAT GTGGAGCGG GGCCGCGGG 60
 GGGCTGCCTG GCCGGTGTG TTGGGGCTG TGCTGGCGCT GTTAGTGCCG GCGGTGGTG 120
 CCGCCAAGAC CGGTGCGGAG CTGCTGACCT GCGGGTCGGT GCTGAAGCTG CTCAATACGC 180
 ACCACCGCGT GCGGCTGCAC TCGCAGACA TCAAATACCG ATCCGGCAGC GGCCAGCAAT 240
 25 CGGTGACCGG CGTAGAGCG TCGGACGAC CCAATAGCTA CTGGGGGATC CCGGGCGGT 300
 CGGAGGCGG GTGCCGCCG GGTCCCCGG TCGCTGCGG GCAGGCGGTG AGGCTCACGC 360
 ATGTGTTTAC GGGCAAGAAC CTGCACACGC ACCACTTCCC GTGGCGGTG TCCAACAACC 420
 AGGAGGTGAG TGCCTTTGGG GAAGACGGCG AGGGCGACGA CCTGACCTA TGGACAGTGC 480
 GCTGCTCTGG ACAGCACTGG GAGCGTGAGG CTGCTGTGCG CTTCAGCAT GTGGGCACCT 540
 30 CTGTGTTCTT GTCAGTCAGG GGTGAGCAGT ATGGAAGCCC CATCCGTGG CAGCATGAGG 600

166

TCCACGGCAT GGCAGTGGC AACAGGCACA ATACGTGGAA GGCATGGA GGCATCTTCA 660
TCAAGCTAG TATGAGGCT TTGGAGGTC ACGATGAAT CTGAGTGTGT GATGGATGG 720
GTGGATGGAG GGTGGAGGT GGGGCTCTG CAGGGCACT CTGGCAGAG ACTTTGGGT 780
TGTAGGGTC CTCAAGTGC TTTTATTA AAGATGTTG GTCTATGAAA AAAAAAAAAA 840
5 AAAAA 844

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 776 base pairs
(E) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TTCGAAATAA AAGATCTGCT CAAGAGAGCC GCAGAAAAG AAGTGTATG TTGGGGGTTT 60
15 AGAGAGCAGG GTCTTAAAT ACACAGCCCA GAATATGAG CTTCAGAACA AAGTACAGCT 120
TCTGGAGGAA CAGAAATTGT CCTTCTAGA TCAACTGAGG AACTCCAGG CCATGCTGAT 180
TGAGATATCA AAAAAACCA GCAGCAGCAG CACCTGCATC TGGTCTTAC TAGTCTCCTT 240
CTGCCTCTC CTGTACTG CTATGTACTC CTCTGACACA AGGGGAGCC TGCCAGCTGA 300
GCATGAGTG TTCTCCGCG AGCTTCGTGC CCTCCGAGT GAGGACCTT ACCAGCTGGA 360
20 GCTGCCTGCC CTGAGTCTG AAGTCCGAA AGACAGTACA CAGCAGTGGT TGGACGGCTC 420
AGACTGTGTA CTCCAGGCTC CTGCAACAC TTCCTGCCTG CTGCATTACA TGCTCAGGC 480
TCCAGTGA GAGCTCCCG TGGAGTGGCC ATTCCTGAC CTCTCTCAG AGCCTCTCTG 540
CCGAGTCCC ATCTCCCGC TGGAGGCAAA TCTCACAAG AAGGAGGAT GGCTTCTTAC 600
25 TGTAGCCCC TGTGTATT TGGAGCAG ATACTCAGGC TAGATATGAG GATATGTGG 660
GGGTCTCAGC AGGAGCTG GGGCTCCCC ATCTGTGTCC AAATAAAAAG CGGTGGGCAA 720
GGCTGGCCG CAGCTCTGT GCTGTCTAG GACGACTGAG GCTCAACA CACCAC 776

(2) INFORMATION FOR SEQ ID NO: 14:

- 30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1376 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```

GAATTCGGCA CGAGGCGCCT ACCCTGCCTG CAGGTGAGCA GTGGTGTGTG AGAGCCAGGC      60
GTCCCTCTGC CTGCCACTC AGTGGCAACA CCCGGGAGCT GTTTTGTCTT TTGTGGAGCC      120
TCAGCAGTTC CCTCTTTCAG AACTCACTGC CAAGAGCCCT GAACAGGAGC CACCATGCAG      180
TGCTTCAGCT TCATTAAGAC CATGATGATC CTCTTCAATT TGCTCATCTT TCTGTGTGGT      240
10 GCAGCCCTGT TGGCAGTGGG CATCTGGGTG TCAATCGATG GGGCATCCTT TCTGAAGATC      300
TTGGGGCCAC TGTCTGCCAG TGCCATGCAG TTGTCAACG TGGGCTACTT CCTCATCGCA      360
GCCGGCGTTG TGGTCTTTGC TCTTGGTTTC CTGGGCTGCT ATGGTGCTAA GACTGAGAGC      420
AAGTGTGCCC TCGTGACGTT CTTCTTCATC CTCCTCCTCA TCTTCATTGC TGAGGTTGCA      480
GCTGCTGTGG TCGCCTTGGT GTACACCACA ATGGCTGAGC ACTTCCTGAC GTTGCTGGTA      540
15 GTGCCTGCCA TCAAGAAAGA TTATGGTTCC CAGGAAGACT TCACTCAAGT GTGGAACACC      600
ACCATGAAAG GGCTCAAGTG CTGTGGCTTC ACCAACTATA CGGATTTTGA GGAATCAGCC      660
TACTTCAAAG AGAACAGTGC CTTTCCCCCA TTCTGTGTGA ATGACAACGT CACCAACACA      720
GCCAATGAAA CCTGCACCAA GCAAAAGGCT CACGACCAAA AAGTAGAGGG TTGCTTCAAT      780
CAGCTTTTGT ATGACATCCG AACTAATGCA GTCACCGTGG GTGGTGTGGC AGCTGGAATT      840
20 GGGGGCCTCG AGCTGGCTGC CATGATTGTG TCCATGTATC TGTACTGCAA TCTACAATAA      900
GTCCACTTCT GCCTCTGCCA CTACTGCTGC CACATGGGAA CTGTGAAGAG GCACCTTGGC      960
AAGCAGCAGT GATTGGGGGA GGGGACAGGA TCTAACAATG TCACTTGGGC CAGAATGGAC      1020
CTGCCCTTTC TGCTCCAGAC TTGGGGCTAG ATAGGGACCA CTCCTTTTAN GCGATGCCTG      1080
ACTTTCCTTC CATTGGTGGG TGGATGGGTG GGGGGCATTC CAGAGCCTCT AAGGTAGCCA      1140
25 GTTCTGTGTC CCATTCCCCC AGTCTATTAA ACCCTTGATA TGCCCCCTAG GCCTAGTGGT      1200
GATCCCAAGT CTCTACTGGG GGATGAGAGA AAGGCATTTT ATAGCTGGG CATAAGTGAA      1260
ATCAGCAGAG CCTCTGGGTG GATGTGTAGA AGGCACCTCA AAATGCATAA ACCTGTTACA      1320
ATGTTTAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAYTCG AGGGGGGTCC CGTACC      1376

```

30 (2) INFORMATION FOR SEQ ID NO: 15:

168

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 502 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

5

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TAAACAGTGG CTTTCCCTCAAG AGGGAGGACT CAGTCAATAT CTGTTGAATG AATGAATGAA 60
 TAATTGCTTG GTTCAACGAA TGAATGGGCTG AATGAATGAT TTCTCCCTTC CTTCCGCACT 120
 GTCTGAGCTG TCCAGGACAG GCAATGGGCGG CAGTGGCTGG TCTGTGGGCT GTCCCACTGG 180
 10 ACTTGGGGTT CTTATGCTTG GTCTGGGGGG AGATCACCCT CAGGCTCCCT AGGTGATCT 240
 TCTGTTCATG GGAATCTGGG TCCGCTCCCTA GTCTTCAGAA CTCACCTGAG GGTGGAGGG 300
 AATACAGGGA GATCTTGGGA GTGCTGAAC AGCGGACAAG AGCGGAGGAG CCGCTGCTTA 360
 AAATGAGGGT GTTGGAGAGG GTTTGGGCTT CTTTMTTGAG TTGAATATGA GATTTCGAG 420
 CAGCCATGAG GATTTGGGTT GGTGGAAGTG GAGATCCGT TCTTCAGTCA GATGGAGGAG 480
 15 GGGGTCCCTT TGGATCTCTT CT 502

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 415 base pairs

20

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

ATCTCTASTG GTGAGTGGCG TGGCTCCAGA CAATGGGAAT CTTGCTTCA CCACCATGG 60
 25 CTGGCTTTCT CTAAGGCTT TGTGGGGGG AGTGGGTTT CAGGAGTTT TTTATCTCT 120
 TGTGATTTT TGGATCAGTG GGAAGACAAG AGGACAGAAG CCAACTTTG TGATTATTT 180
 GGCCGATGAC ATGGGGTGGG GTGACTGGGG AGCAACTGG GCAGAAACAA AGGACACTG 240
 CAACCTTGAT AAGATGCTT CAGAGGGAAT GATGTGATC TTGATATGCC AGCCAGCTT 300
 TCTTTGAAAG TCTTACTCCC GTTCTTAAA AGGGAAGGG GGTGCAAG CACTTAAGA 360
 30 WTGATGATG GATCCATGTG ATTTATTTAA TTTATTAATT AATTGGITT GGAACCCAG 420

ATAGC

425

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1316 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

10 GGCACGAGGA GCTGGGGGAG CCGAGGTCG GCTACGTGGC TGGCATGCAT GGGAACGAGG 60
CCCTGGGGCG GGAGTTGCTT CTGCTCCTGA TGCAGTTCCT GTGCCATGAG TTCCTGCGAG 120
GGAACCCACG GGTGACCCGG CTGCTCTCTG AGATGCGCAT TCACCTGCTG CCCTCCATGA 180
ACCTGATGG CTATGAGATC GCCTACCACC GGGGTTGAGA GCTGGTGGGC TGGGCCGAGG 240
GCCGCTG3AA CAACCAGAGC ATCGATCTTA ACCATAATTT TGCTGACCTC AACACACCAC 300
15 TGTG3SAAGC ACAGGACGAT GGAAGGTGC CCCACATCGT CCCCACCAT CACCTGCCAT 360
TGCCCACTTA CTACACCTG CCCAATGCCA CCGTGGCTCC TGAAACGCGG GCAGTAATCA 420
AGTGGATGAA GCGGATCCCC TTGTGCTAA GTGCCAACCT CCACGGGGT GAGCTCGTGG 480
TGTCTACCC ATTGACATG ACTGCGACCC CGTGGGCTGC CCGCGAGCTC ACCCCACAC 540
CAGATGATGC TGTGTTTCGC TGGCTCAGCA CTGTCTATGC TGGCAGTAAT CTGGCCATGC 600
20 AGGACACCAG CCGCCGACCC TGCCACAGCC AGGACTTCTC CGTGCACGGC AACATCATCA 660
ACGGGGCTGA CTGGCACACG GTCCCCG3GA GCATGAATGA CTTAGCTAC CTACACACCA 720
ACTGCTTTGA GGTCACTGTG GAGCTGTCTT GTGACAAGTT CCCTCAGGAG AATGAATTGC 780
CCCAGGAGTG GGAGAACAAAC AAAGACGCC TCCTCACCTA CCTGGAGCAG GTGGGCATGG 840
GCATTGCAGG AGTGGTGAGG GACAAGGACA CCGAGCTTGG GATTGCTGAC GCTGTCAATG 900
25 CCGTGGATGG GATTAACCAT GACGTGACCA CGCGTGGGG CGGGATTAT TGGGCTCTGC 960
TGACCCACAG GGAATACATG GTGACTGCCA GTGCCGAGGG CTACCATCA GTGACACGGA 1020
ACTGTGGGGT CACCTTTGAA GAGGGCCCTT TCCCTGCAA TTTCGTGCTC ACCAAGACTC 1080
CCAAACAGAG GCTGCGCGAG CTGCTGGCAG CTG3GGCCAA GGTGCCCCCG GACCTTCGCA 1140
GGGCGCTGGA GCGGCTAAGG GGACAGAAGG ATTGATACCT GCGGTTTAAG AGCCCTAGGG 1200
30 CAGGCTGGAC CTGTCAAGAC GGAAGG3GA AGAGTAGAGA GGGAGGGACA AAGTGAGGAA 1260

AAGGTGCTCA TTAAAGCTAC CGGTACCTT AAAAAAAAAA AAAAAAAAAA AAAAAA 1316

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 436 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

10 AAAAAATTC AATGGATATT ATGAAATTA GAGAGTATTT CCAGAAGTAT GATATAGTC 60
CACGTGTCAA GAAAAATTC GTACACGAGC AAGAAGCCAT TAACTCTGAC CCAGAGTTGT 120
CTAATTGTGA AAATTTTCAG AAGACTGATG TGAAAGATGA TCTGTCTGAT CCTCCTGTTG 180
CAAGCAGTTG TATTCTGTAG AAGTCTCCAC GTAGTCCACA ACTTTCAGAT TTTGGACTTG 240
AGCGGTACAT CGTATCCCAA GTTCTACCAA ACCCTCCACA GGCAGTGAAC AACTATAAGG 300
15 AAGAGCCCGT AATTGTAAOC CCACCTACCA AACAATCACT AGTAAAAGTA CTAAAAACTC 360
CAAAATGTGC ACTAAATGGS ATGATTTTGA GTGIGTACTC CTAAATTAGA AACTTTGGT 420
ATCTCTGAAT ATACTA 436

(2) INFORMATION FOR SEQ ID NO: 19:

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 503 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

TGTGCATATC CTGGGAAAA AAATGGTACA TGTTTTAGAA ATTTTACTGT TTATAACAAT 60
GCAGGCAGTC AGTTTCCCGT TTCAAACACA GATAGATACA TGCAACACTC AAGATCCTGC 120
AGAGAGGCAG CCAGCATCTA TTGTTTAAAA AGGTTTCAAA AAGAATTCGG ATGCTCKTT 180
TCTCTTTTGA ATCTGTGTGC CAAATGACAG GACCAATAT TCGTCTCTT TTCKGTAA 240
30 AYTCAAGAAAG AMACATGAAA GAACCCAGAA TGCATTTCTT AAAGGGATTY AGTGCAGTTA 300

171

TTTTAAATAA TTTATGCAGC CACACACACA TACATATATC CCCCAGTAC ATATTTTTC 360
 CCTTTTACT TGTGTGCAAT CAGTAGCTAC AATGACTGAA ATCCACTTCT TTGGGACTGT 420
 GACATTTAAG CAAATCTTGT NTCTAGAAAN CGAAATGCCA NANTCTCGCA CAAAGCTGCT 480
 CCGTCTGGGG CAACAAATTC ACA 500

5

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 358 base pairs

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

GGGCTGTCTC CCCAGTAGTA ACTTGCTGGC CCTGCCCTTG AAGTGGGGAA ACTGTGAAGG 60
 GCTCCTTGAT CAAGCTTGTG CTCTTTTCTT ACCTCTTCCT CTCCTCTGTT TCCGCTGCAG 120
 15 CTGAACAGGC CAGCAGGCAA CCTGCCATGG GGTCTGCTC CAAGAACCGG TCCTTCTTCT 180
 GGATGACTGG GCTCCTGCTA TTCATCAGCC TCCTCCTCAG TGAGTGGCAG GGTCCCTGGG 240
 AAGGGAGGGC AATTGGAGAG GGCTGGGCTA GCTGGGCTCT GACCAACGGG TGGGCTGTTC 300
 AACTTCTGAT GTCTTTGGGC AACAAACAG AAAAACACTC TGTATGATT TACGAAAN 358

20

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1926 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

25

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

AGTGAAGGGA GCTGGCCGTG CAGCTGGGCT TCGGGCCCTG TGCCAGAGGA GCAGCCCTTC 60
 CTGAGCAGGA GGAAGCAGGT GGTGGCCGCG GCCTTGAGGC AGGCCCTGCA GCTGATGGA 120
 GACCTGCAGG AGGATGAGAT CCACTGCTTA GCTATTATGG CCACTGGTGG TGGGATCCGG 180
 30 GCAATGACTT CCCTGTATGG GAGCTGGCT GGCCTGAAGG AGCTGGGCCT CTTGATGTC 240

	KTCTCTTACA TCACCTGGAC CTCTGGCTCC AGCTGGGCTT TACCTAACCT TTATAAGGAC	300
	CCAGATGGT CTGAAAGAA CTGGGCAGGG CCATCTGAST CTTTAAGAC CCAGGTGACC	360
	AAGAAACAGC TGTATGTTT GACCTCCAGC CAGCTACAGC GTTAAGGCA GSACTGGCC	420
	GAGCTGCCC GTTAAAGTA CCAAGCTGC TTAAACAAC CTGGAGGCT CATCAACGAG	480
5	GGGTGCTGC ATGACTAGTC CCATGATCAC AAGCTCTCAG ATCAACGGA GGCCTGAGT	540
	CATGGCCAGA ACGTTCTGC CATCTACTGT GCTCTAACCA CTAAGGGA GAGCTGACC	600
	ACTTTTAAAT TTGGGAGTG GTGGAGTTC TCTCCCTAGC AGGTGGGTT CCCCAGTAC	660
	GGGCTCTCA TCTCTCTGA GCTCTTGGC TCCGATTCT TATGGGCA GCTGATGAAG	720
	AGGCTTCTG AGTCTCGAT CTGCTCTTA GAAGTATCT GAGTAACCT GTATGCAGCC	780
10	AACCTCCAGG ACGTTTATA CTGGCTCA GAGCCAGTC AGTTTGA GCGTGGGTG	840
	AGGAACAGG CCAACTGGA CAAGGAGCAG GTCCCCCTC TGAATAGA AGAACCACC	900
	TCAACAGCCG GCAATATAG TGAATTTTC ACCGATCTC TGAAGTGGC TCCACTGGCC	960
	CAGGCCACAC ATAACTTCT GGTGGGCTC CATTCCACA AAGACTACT TCAGCATCCT	1020
	CACTTCTCCA CATGAAAGC TACCCTGTG GATGGCTCC CCAACAGCT GACACCTCG	1080
15	GAGCCCCACC TGTGCTGCT GATGTGTC TACCTCATCA ATACAGCTG CCTGCCCCCT	1140
	CTGAGCCCA CTGGAGCT GAGCTCATC CTGTATTGG ATACAACT CCACGGAGCC	1200
	TTCCAGCAGT TGAAGTCTT GGGCCGGTC TGCCAGGAG AAGGATCCC GTTCCCACC	1260
	ATCTCGCCA GCGCGAAGA GCAGCTCCAG CCTCGGAGT GATACACTT CTCCGACCC	1320
	ACCTGCCCC GAGCTCTCT GTGCTGCAC TTCTCTTGG TGAAGATC CTTCGGGAG	1380
20	TACTCGCCCC CTGGGCTCC GGGACATCC GAGGAGGCG CAGTGGGA GGTGAACCTG	1440
	TCTTCATCG ACTCTCCTA CCACTACAG AAGGTGACCT ACAGTCAGGA GGACGTGGAC	1500
	AAGCTGCTGC ACCTGACACA TTAAATGTC TGCAACAAC AAGAGCAGCT GCTGGAGGCT	1560
	CTGCGCCAGG CAGTGAGGS GAGGCGGAG CCGAGGCCC ATGATGACC GGGCCCCCTG	1620
	CCACCCCTAA CTCTCATCA TTCTCTGCT GTGAGTTGC AAGTGGAAAC TGTATCACG	1680
25	CAGTGCTTNC AGAGCTGCG GTTCAGCTGG CACTGTCCA GGTTCAGGC TGAGGGCTG	1740
	GAGCTCCCTT GCGCTCAGC AGTTGCACT GGGTAAGGA GGTCAAGCC ATTTGTGTAA	1800
30	TCACCCAAAA CCCCCGGGC TGTGCTGT TTCCCTCTG CGTACCTG AGTAGTTGGA	1860
	GCACTTGATA CATACAGAC TCATACAAAT GTGAGGCGCT GAGAAAAAA AAAAAAAA	1920
	ACTCGA	1926

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1224 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

5	CCGCCGAAGC TCCGTCCCGC CCGCGGCCGG CTCGCCTCA CCTCCCGGCC GCGGCTGCCC	60
15	TCTGCCCCGG TTGTCCAAGA TGSAGGGGCG TCCACCGGGG TCGCTCGCCC TCCGGCTCCT	120
	GTTGTTCTGT GCGCTACCGG CCTCCGGCTG GTTGACGACG GCGCGCCCCG AGCCGCCGCC	180
20	GCTGTCCGGA GCCCCACAGG ACGGCATCAG AATTAATGTA ACTACACTGA AAGATGATGG	240
	GGACATATCT AAACAGCAGG TTGTCTTTAA CATAACCTAT GAGAGTGGAC AGGTGTATGT	300
	AAATGACTTA CCTGTAAATA GTGGTGTAA CCGAATAAGC TGTGAGACTT TGATAGTGAA	360
25	GAATGAAAAT CTTGAAAATT TGSAGGAAAA AGAATATTTT GGAATTGTCA GTGTAAGGAT	420
	TTTAGTTTCAT GAGTGGGCTA TGACATCTGG TTCCAGTTTG CAACTAATTG TCATTCAAGA	480
30	AGAGGTAGTA GAGATTGATG GAAAACAAGT TCAGCAAAAG GATGTCACTG AAATTGATAT	540
	TTTAGTTAAG AACCGGGGAG TACTCAGACA TTCAAACCTAT ACCCTCCCTT TGGAGAAAAG	600
35	CATGCTCTAC TCTATTTCTC GAGACAGTGA CATTTTATTT ACCCTTCCTA ACCTCTCCAA	660
	AAAAGAAAGT GTTAGTTTAC TGCAAAACCAC TAGCCAGTAT CTTATCAGGA ATGTGGAAAC	720
	CACTGTAGAT GAAGATGTTT TACCTGGGCA AGTTACCTGA AACTCCTCTC AGAGCAGAGC	780
40	CGCCATCTTC ATATAAGETA ATGTGTCACT GATGGAAAA GTTTAGAAAA GATCTGTGTA	840
	GGTTCTGGAG CAACGTTTTG CCAGTATTCT TTCAGTTTTT GAACATCATG GTGGTTGGAA	900
45	TTACAGGAGC AGGTGTGATA ATAACCATCT TAAAGGTGTT TTTCCCAGTT TCTGAATACA	960
	AAGGAATTCT TCAGTTGGAT AAAGTGGACG TCATACCTGT GACAGCTATC AACTTATATC	1020
	CAGATGGTCC AGAGAAAAGA GCTGAAAACC TTGAAGATAA AACATGTATT TAAAACGCCA	1080
50	TCTCATATCA TGACTCCGA AGTAGCCTGT TGCCTCCAAA TTGCCCCTT GAATATAATT	1140
	TTCTTTAAAT CGTTAAGAAT CAGTTTATAC ACTAGAGAAA TTGCTAAACT CTAAGACTGC	1200
55	CTGAAAATTG ACCTTTACAG TGCC	1224

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 694 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

10 GGCACGAGTC TTATTGTGCA CTGTAGCCTG AATCCCCCAG GGTAAATTAAT ATGAAGTGCA 60
 AAAAGTTGAA TTTTCAGTC TAAAAGGCAG TGGGAGAAAT TACATAGCAT GGAAATAATA 120
 15 AAATGAAGTC TTATTAATGA GAACGAGGCT CTGTGAGTGG CAAGTTCTGC TGSTCACCCG 180
 ATGGGATGAG GAGGCTTTCA AGCTTTTTTT TGGTAATAC TCACAGTTTC CAACGTCTGT 240
 GTACTTTTCA AAATGAGCTT GTTCTTCCTT CTGACACACA TGTCAAAGCT CCATGGTGAC 300
 20 GCAGAGGTCT GTTAAGGTC ACAGTCTCTC GTTGTGATTG GCATACGGTC CTGTAGCAT 360
 ACTTGTTAGC CCACTGCTGC TTGAAGGAAC TAAGAGTATT CAGGGATAGA GAGCTGAAAA 420
 25 TAGGATTAAT TCTTCTCTTT TGAATCTCCC CTCAAGATGT CCTTGCTTTG GTCTGAAAAA 480
 CTCTCTGAC AACTTTTCCC CAAAGCAAAC CATCTGCTTT TGTGAACTC TGAGTGAATA 540
 TATTAGATC TTCTCTCTG AGCCTCTGTA CTGTCANSTT TGTCTGTTG TTTGTTTCCA 600
 30 AGAGACTGTG TCTTCTCTG TCACCCAGGA GTTTGAAACC AGCTTGCAA CATAGCAAG 660
 CCTATCTCT ACAAAAAAAA AAAAAAAA AAAA 694

35

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 796 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
- 45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

ATGAGGAGCG GTTGTATGCG GCAGGTGGA GTTGTGTGAA CAGGGGTCTT GGGCTGCG 60
 50 CTGTGTGTCG TGTGTGCTT CCGACTAGCG CTGAGGCGCG CCGAGTCCCG CTTTCCACCT 120
 CGACCTCTGC CAGGCGCGCA CCGAGCTCA GGTGTGTGCG CAGCCACCAA GTTCCAGTG 180
 55 CGCACCAGTG GTTATGTGCT GGCCTCACC TGTGTGTGCG ACAGGACTTG GACTGCAGCG 240
 ATGGCAGCGA TGAGGAGGAG TGCAGGATG AGGCATGTAC CCAGAAAGGG CAATGCCAC 300
 CGCCCCCTTG CTTCCCTTGC CCGTGCACCG GGTGAGTGA CTGTCTGGG GGAAGTACA 360
 60 AGAAACTGCG CAACTGCAGC CCGCTGGCCT GCTAGCAGS GRAGSKCMCG WKGCACGCTG 420

AGCGATGACT GATTTCACCT CACGTGGCGG TGGGACGGCC ACCGAGACTG TCCCGACTCC 480
 AGCGACGAGC TGGGTGTGG AACCAATGAG ATCTCTCCGG AAGGGGATGC CACAACCATG 540
 5 GGGCCCCCTG TGAACCTGSA GAGTGTACCC TCTCTCAGGA ATGTCACAAAC CATGGGGCCC 600
 CCCTGACCC TGGAGAGTGT CCCCTCTGTC GGGAAATGCCA CATCTCTCTC TGCCGAGAG 660
 10 CAGTCTGSA GGGCAACTGC CTATGGGGTT ATTGCAAGCTG CTGCGGTGCT CAGTGTAACT 720
 CTGCTACCG CACCCCTCTT CCTTTTCTCC TGGTCCGAG CCCAGGAGCG CCTCCGCCA 780
 CTGGGGTTAC TGGTGG 796
 15

20 (2) INFORMATION FOR SEQ ID NO: 25:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 662 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

30 TAATTGGCA CAGGCTGTG GTGGAGAAGG ACGTCCGTG CCGCTGGGT CTGAGCCGGA 60
 GTGGTGGTG GGTGGATGG AGGCGACCTT GAGCAGCAC TTGGAAGACA CAATGAAGAA 120
 TCCCTCCATT GTTGGATCC TGTGCACAGA TTCACAAGGA CTTAATCTGG GTTGGCGCG 180
 35 GACCCTGTCA GATGAGCATG CTGGAGTGAT ATCTTTCTA GCCCAGCAAG CAGCTAAGCT 240
 AACCTCTGAC CCGACTGATA TTCTGTGTGT GGTCTAGAA TCAGATAATG GGAACATTAT 300
 GATCCAGAAA CAGGATGGCA TCACGGTGGC AGTGACAAA ATGGCCTCTT GATGCTCATA 360
 40 TCTGTCTTC AGCAGCTGT CATAGGAAC TGGTCTACC TATGTTAATT ACCTTATAGA 420
 ACTACTAAG TCCAGTAGT TAGGCCATTC ATTAAATGTG CATTAGGCAC TTTTCTGTT 480
 45 ATTTAAGAGT CAATTGCTTT CTAATGCTCT ATGGACCGAC TATCAAGATA TTAGTAAGAA 540
 AGGATCATGT TTTGAAGCAG CAGGTCCAGG TCACTTTGTA TATAGAATTT TGCTGTATT 600
 AATAAATCTG TTGGAGGAA AAAAAAAAAA AAAAAAATTA CTGCGGNCCG ACAAGGGAA 660
 50 TC 662

55

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1105 base pairs
 60 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26

5 CCTGATCTTC TCTTTTCTGC AGTTCAAGGG AAAGACGAGA TCTTGACAAA GGCACCTCTGC 60
TCTGTCCTTT GGCCTGGGAA GGGTGGGATG GAGCTCTCC GCTGCTCAT CTTACTCTTT 120
10 GTACAGAGC TGTCCGGAGC CCACAACAC AATSTSTTCC AGGGGTGGC GGGCCAGTCC 180
CTGAGGTST CTGCCCCCTA TGACTCATG AAGCACTGGG GAGGCGCAA GGCCTGCTG 240
CGCCAGTTGG GAGAGAAGGG CCCATGTCAG CTTSTSTCA GCACGCACAA CTTSTSTCTG 300
15 CTSTCTTCC TGAGGAGTG GAATGGGAGC ACAGCCATCA CAGACGATAC CTTGGGTGGC 360
ACTCTACCA TTACGCTGCG GAATCTACAA CCCCATGATG CGGTCTCTA CCACTGCCAG 420
20 AGCTTCATG GAGTGAGGC TGACACCTC AGGAAGGTCC TGSTGAGST GTTCGCAGAC 480
CCCCTGATC ACCGGGATGC TGGAGATCTC TGGTCCCCG GGGAGTCTGA GAGCTTCGAG 540
GATGCCCATG TGGAGCACAG CATCTCCAGG AGCTCTTCT AGGAAAGGCC GCAATTCCT 600
25 ATTCTTCCC CTCTTGCTA TCTTCTCT CCAAGAYCTG CATCTTCTC ATCAAGATT 660
TAGCAGCCAG CGCCCTCTGG GCTGCAGCT GGCATGGACA GAAGCCAGGG ACACATCCAC 720
30 CCACTGAAT GACTGTGGC CATGACCCAG GGTATCAGCT CCAAACTCTG CCAGGGCTGA 780
GAGACACCTG AAGGAAGATG ATGGGAGSAA AAGCCAGGA GAAGTCCAC CAGGGACCAG 840
CCCAGCCTGC ATACTTGCCA CTGGCCACC AGGACTCCTT GTTCTGCTCT GGCAAGAGAC 900
35 TACTCTGCT GAACACTGCT TCTCTGGAC CTTGGAAGCA GGGACTGTT GAGGGAGTGG 960
GGAGCTGTA AGAACACCTG ACACTTCTG AATATTGGAC ATTTTAAACA CTTACAAATA 1020
40 AATCCAAGAC TGTATATTT AAAAAAAAAA AAAAAAAMA AAARFRRRC CCCGTACCC 1080
AATTGCCCC ATAGTGAGTC GTATA 1105

45

(2) INFORMATION FOR SEQ ID NO: 27

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 1017 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

CTCGCTGGG CTGTTTCCCG GCTTCATTTC TCCGACTCA GCTTCCACC CTGGCTTTC 60
CGAGGTGTT TCGCCGCTGT CCCACCACT GCAGCATGA TCTCTTAAC GGACACGCAG 120
60

AAAATTTGGAA TGGGATTAAC AGGATTTGGA GTSTTTTTC TGTCTTTGG AATGATTCTC 180
 TTTTTTGACA AAGCACTACT GGTATTTGGA AATGTTTAT TTGTAGCCGG CTGGGCTTTT 240
 5 GTAATTGGTT TAGAAAGAAC ATTGAGATTC TTCTTCCAAA AACATAAAAT GAAAGCTACA 300
 GGTTTTTTTC TGGGTGGTGT ATTTGTAGTC CTATTTGGTT GGCTTTTGAT AGGCATGATC 360
 10 TTGAAATTT ATGGATTTTT TCTCTGTTT AGGGGCTTCT TTCTGTCTGT TGTGGCTTTT 420
 ATTAGAAGAG TGCCAGTCCT TGATTCCTC CTAATTTTAC CTGGAATTAG ATCATTGTGA 480
 GATAAAGTTG GAGAAAGCAA CAATATGCTA TAACAACAAG TGAATTTGAA GACTCATTTA 540
 15 AAATATTTGT TTATTTATAA AGTCATTTGA AGAATATTCA GCACAAAATT AAATTACATG 600
 AAATAGCTTG TAATGTTCTT TACAGGATTT TAAACGTAT AGCCTACAAA GTACCAGCAG 660
 CAAATTAGCA AAGAAGCAGT GAAACAGGC TTCTACTCAA GTGAACTAAG AAGAAGTCAG 720
 20 CAAGCAAACCT GAGAGAGGTG AAATCCATGT TAATGATGCT TAAGAACTC TTGAAGGCTA 780
 TTTGTGTGT TTTTCCACAA TGTGCGAAAC TCAGCCATCC TTAGAGAACT GTGGTGCCTG 840
 25 TTCTTTTCT TTTTATTTTG AAGGTCAGG AGCATCCATA GGCATTGTGT TTTTAGAAAT 900
 GTCCACTGCA ATGGCAAAAA TATTTCCAGT TGCATGTAT CTCTGGAAGT GATGCATGAA 960
 30 TTCGATTGGA TTGTGTCATT TTAAATATT AAAACCAAGG GAAACCCCAA AAAAAA 1020

(2) INFORMATION FOR SEQ ID NO: 28:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 391 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

45 CCCTGGAAAG AGGAACTGAT GTTTGAGGGG ACAGATGTGG GTCATTTCCT CTGGCAGTGC 60
 CCTCTAGCCT TGCTGCCTTG GCTTTCTGAC CCTTCCAGG CTCAGGGGC CTGGGAGATC 120
 TCATGCCTCA GCCCAGGAAA CATTTAATAG GGAAAGCAGA GACATGTCAT GTCAGCCCCA 180
 50 CAGACAAGAA TTCTAGAGC ACTTGTCCTG TTGTTCTTTG CCCCACATT ACTCAGTCTG 240
 GGCCATGGAA TCCATCCAAT AAACACAGCA ACACCCTATG NTAATGACCA AGCAAAGCTT 300
 GCCCCTGGTA CCAAAGAGCT AAATCATGAC CAAAGTGTGA CATGAATGTA ACTGAAATGC 360
 55 GGTTAGTTG CTCAATGTAT GCAAAGTCCC A 391

60

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1139 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

GGTGATATCT TCATAGTGGG CTATTACAGG CAGGAAAATG TTTTAACTGG TTTACAAAAT 60
CCATCAATAC TTGTGTCATT CCTGTAAAA GGCAGGAGAC ATGTGATTAT GATCAGGAAA 120
CTGCACAAAA TTATTGTTTT CAGCCCCCGT GTTATTGTCC TTTTGAACGT TTTTTTTTTT 180
ATTAAAGCCA AATTGAGGT CTATATATTC GTATTCCATG TGTTAGATGG AAGCATTTCC 240
TATCCAGTGT GAATAAAAG AACAGTTGTA GTAAATTATT ATAAAGCCGA TGATATTTCA 300
TGGCAGGTTA TTCTACCAAG CTGTGCTTGT TGGTPTTTCC CATGACTGTA TTGCTTTTAT 360
AAATGTACAA ATAGTTACTG AAATGACGAG ACCCTTGTTT GCACAGCATT AATAAGAACC 420
TTGATAAGAA CCATATTCTG TTGACAGCCA GCTCACAGTT TCTTGCTGA AGCTTGGTGC 480
ACCCTCCAGT GAGACACAAG ATCTCTCTTT TACCAGGT GAGAACAGAG CTGGTGGATT 540
AATTAATAGT CTTGATATC TGCCATGGG TAACCTCATT GTAACATCA TCAGAAATGGG 600
CAGAGATGAT CTTGAAGTGT CACATACACT AAAGTCCAAA CATTATGTCA GATGGGGGTA 660
AAATCCATTA AAGAACAGGA AAAAATAATT ATAAGATGAT AAGCAAATGT TTCAGCCCAA 720
TGTCAACCCA GTTAAAAAAA AAATTAATGC TGTGTAAAT GATTGAATTA GTTTGCAAAC 780
TATATAAAGA CATATGCAAT AAAAGTCTG TTAATGCACA TCTGTGGGA ATGGAGTGT 840
CTAACCAATT GCTTTTCTT GTTATCTGAG CTCTCTATA TTATCATACT CAGATAACCA 900
AATTAAAAGA ATTAGAATAT GATTTTAAAT AACTTAAACA TTAAACTCTT CTAACTTTCT 960
TCTTCTGTG ATAATTGAGA AGATAATTAT GGATCTTCAA TGCTCTGAG TCATTGTTAT 1020
AAAAATCAG TTATCACTAT ACCATGCTAT AGGAGACTGG GCAAAACCTG TACAATGACA 1080
ACCCTGGAAG TTGCTTTTTT TAAAAAATA ATAAATTTCT TAAATCAAAA AAAAAAAA 1139

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 465 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCACGCGTCC GCGGACGCGT GGGGAAGGTT TGTGCCAGTA GACATTATGT TACTAAATCA 60
 5 GCACTTTAAA ATCTTTGGTT CTCTAATTCA TATGAATTG CTGTTTGCTC TAATTTCTTT 120
 GGGCTCTTCT AATTGAGTG GAGTACAATT TTGTTGTGAA ACAGTCCAGT GAAACTGTGC 180
 AGGGAAATGA AGGTAGAATT TTGGGAGGTA ATAATGATGT GAAACATAAA GATTTAATAA 240
 10 TTACTGTCCA ACACAGTGGA GCAGCTTGTC CACAAATATA GTAATTACTA TTTATTGCTC 300
 TAAGGAAGAT TAAAAAAGA TAGGGAAAAG GGGGAACTT CTTTGAAAAA TGAAACATCT 360
 GTTACATIAA TGTCTAATTA TAAAATTITA ATCCTTACTG CATTTCTTCT GTTCTACAA 420
 15 ATGTATTAAA CATTCAGTTT AACTGGTAAA AAAAAAAAAA AAAAA 465

20

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 702 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

30 GCAACAAGCG GCCACCTTC CTGAAGATCA AGAAGCCACT GTCGTACCGC AAGCCCATGG 60
 ACACGGACCT GGTGTACATC GAGAAGTCGC CCAACTACTG CGAGGAGGAC CCGGTGACCG 120
 35 GCACTGTGGG CACCCAGGGC CGCGCCTGCA ACAAGACGGC TCCCCAGGCC AGCGGCTGTG 180
 ACCTCATGTG CTGTGGGCGT GGCTACAACA CCCACCACTA CCCCCGCGTG TGGCAGTGCA 240
 40 ACTGTAAGTT CCACTGGTGC TGCTATGTCA AGTGCAACAC GTGCAGCGAG CGCACGGANG 300
 ATGTACACGT GCAAGTGAGC CCCGTGTGCA CACCACCCTC CCGCTGCAAG TCAGATTGCT 360
 GGGAGGACTG GACCGTTTCC AAGCTGCGGG CTCCCTGGCA GATGCTGAG CTGTCTTTT 420
 45 CTGCTGAGGA GGGTACTTTT CCTGGGTTTC CTGCAGGCAT CCGTGGGGGA AAAAAAATCT 480
 CTCAGAGNCC TCAACTATTC TGTTCCACAC CCAATGCTGS TTCACCCTCC CCCAGACACA 540
 50 GCCCAGGTCC CTCGCGGGCT GGAGCGAAGC CTTCTGCAGC AAGAACTCTG GACCCCTGGC 600
 CCTCATCACA GCAATATTTA ACAATTTATT CTTGATAAAA ATAATATTAA TTTATTTAA 660
 TAAAAAGAAT TCTTCCAAAA AAAAAAAAAA AAAAAAACNT CG 702

55

(2) INFORMATION FOR SEQ ID NO: 32:

60

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1141 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

	CGGTACGAGG AAGAAATGGC AGAGACTGGA ATCTCTCTTC ATGAAAAAAT GGAGCCCTT	60
10	AACCTCAGTT CGACAGAGTG CAGTTCCTTC TCTCCACCCA CCACAGTGAT TCTCCTTATC	120
	CTGCTGTGCT TTGAGGGGCT GCTCTTCTTC ATTTTCACAT CAGTGATGTT TGGGACCCAG	180
	GTGCACTCCA TCTGCACAGA TGAGACGGGA ATAGAACAAT TGAAAAAGGA AGAGAGAAGA	240
15	TGAGCTAAAA AAGCAAAATG GATGAACATG AAAAGCGTTC TTGGCCACCC CTCTCTCTA	300
	GGTTGGGCGA GCGGCTTTGC CAGGTCAGAG CAAGGGAAGG CAGACCCGTA CCAGTATGTC	360
20	GTGTGAAGAA CCGGACCCGG CATGCGCACT CAGACACAAG TCCACACCAC AGCACTACCG	420
	TCCCATCGCT TCTCATGAAT GTTAAATCG AAAAAGCAAA ACACTACTC TTAATACTTC	480
	TTTATCTCT CAATAAAAAT GGCTGAGCAT TGCAGAGARA AAAAAAAGTC CCCACATTTC	540
25	ATTTTAAAA AACCATCTTC TCGATTCTTC TTGCTGACCG AAGCTGCTCT CTTTCTCTTC	600
	TAAAACTACT TCTCTGCGCT CTGCTTTCTC TCTGCTGTCT GTCTGGCATG ACTAATGTAG	660
30	AGGCGGCTCT CTGCGGCTGT GCGCATCTTA CTAAGTGAGT GAGACATGAC GTGTGCTGG	720
	GATGGAAATG TCTGGACACC TGCTGGGGA TGCATGGAAG AGCCAGGAGG GTCTGACCT	780
	TCCCACTGCT CAGGAGGAG TGCTGGGCTC CCGATGGA CATAAACCT CACCGAAGAT	840
35	GGATGCTTAC CCCTTGAGGC CTGAGAAGCG CAGGATCAGA AGGACCTTC GCACAGCGAC	900
	CTCATCCCCC AAGTGGACAC GCTTTCCTTC CTAAGTCGCA AAGCAATTGC CTGCTTCTA	960
40	CTTTATGGG TTGGGCTGTG TACAATGATT TTGCGGGGGA GTGGGGGAGA AAGATGAAAG	1020
	AGCTCTTATT TCTATTCTGA ATCAGCAATT ATATCCCTG TGATATTTG GAAGACTGTC	1080
	TAGGAAGAGC GTTTTCCAG TTCAAAATGC CTTATACAAT CAAGAGGAAA AAAAAAAAAA	1140
45	AG	1142

50

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

55

(A) LENGTH: 928 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

60

GGCACGAGGT CTAATGAGGG CTCTCTTGTG TGCTAGAGAT GAGAGAAATG TATACTAATC 60
ATTCTAATTT GTACTTAAAA TACATTTTAC TAATCATATT GATTTTAAAT ATGACAAATT 120
5 CTCTTAGTAG ATACTAATCT TTCTTGTTTA TCATATTGTC CTAGAGAAGC CTAGSTAAAA 180
ATGGSTTCCA CCTAGTCTGT TTGTATAACA CCTTCCCCCG TCCCCTCTCC ATCCCTGCCA 240
ATTGGGCTCT ATGCATATTG ACAAGCAAAT AAGAAAACCT TAGGTTCTTG TATTIGAATT 300
10 TCCAAAACAA TAAAAGGTTT TGAICTAAGA TTGCAATTCA AGAAGAGGCA GAAATTTTGT 360
CTTATCTTCT TATCATTTTG TGAAGTTGTG TTTCTCTGTA TGCTTAGAAA ATTTACACAC 420
15 AAGGAATGTT TGAAAAAGTG AGAATTTTAG AGTGCTTGGG TGGTTTITAT TTGGTCAGTG 480
CTGATCTGTT AGGTGTTTAG GGAATAATG CTTCAGGACC TTTTGTACAA CACAGCTTCA 540
TGAATGACTG GGGATATTT ATGTTTGTGC TGAGAAAAGG GAGGGAGTGG GCAGGTTGGA 600
20 GTGGGACCT TTCCATTGAA AGCAGTGCG AGCTGTTTT CTAGATGCA TTTTCTCTT 660
ATGCTTGTA CATGTTCTT GTGTCCATAA TTGACTGAAA TGTCAAGCTC CAGGAATGCA 720
25 AGGCATTIAT CAGGTGACCA GAAGTAGAAC CTTGTTGATT ATGAAATGGA AGAATAATGT 780
CAAGGTACTG GGGGTAAAAT GACAAATAAG ATTTTACTGG TGAATTTCCA TGCTTAGTAT 840
GTACATTAA CTCTTTTAA GTTGCATGTT AATCTGTTAT AACGTATTGT GTCTGGTTTA 900
30 TGCTTTGAGT AAAAAAAAAA AAAAAAAA 928

35

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

40

- (A) LENGTH: 773 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

45

GGCACGAGTT CTGGCTCTC ATTCTCTTAC ACTCTGACAT GAATGAATTA TTATTATTTT 60
TCTTTTCTT TTTTITTTT ACATTTTGTA TAGAAACAAA TTCATTTAAA CAAACTIAT 120
50 ATTATTATTT TTTACAAAAT ATATATATGG AGATGCTCCC TCCCCCTGTG AACCCCCCAG 180
TGCCCCCGTG GGGCTGAGTC TGTGGGCCCA TTCGGCCAAG CTGGATTCTG TGTACCTAGT 240
ACACAGGCAT GACTGGGATC CCGTGTACCG AGTACACGAC CCAGGTATGT ACCAAGTAGG 300
55 CACCTTGGG CGCACCCACT GGGGCCAGG GTGGGGGAT GTTGGGAGCC TCCTCCCCAC 360
CCACCTCCC TCACTTCACT GCATTCCAGA TTGGATATGT TCCATAGCCT TGCTGGGGAA 420
60 GGGCCCACTG CCAACTCCCT CTGCCCCAGC CCCACCTTG GCCATCTCCC TTTGGGAAC 480

AGGGGGCTGC TGGTGGGAAA TGGGAGGTAG GGCAGATGTA TGCATTCCCTT TATGTCCCTG 540
 TAAATGTGGG ATTACAAGAA GAGGAGGTGC CTGATGTGTA CTTCTCTCTC CTGTAATCC 600
 5 TGTGGGCGAG CTTTATGGCA GAAATAGGT ATTTTATGGC TATTTTGTGTA ATAGGGCTTC 660
 TGGTCAAAAT CCTGTGTAG CTGAATTCC AGGCTCTGCA TTGTACAGCC CCCCCTCCC 720
 10 CTCACCACT AATAAGGAA TASTTAACAC TAAAAAATA AAAAAAATA AAA 775

15 (2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

25 TAAAAATGTA CAGCCTGTG ATATTCCAGG CACTGCACTA TGTATGCCGT TTATCAACAG 60
 TTAGCTCAGC TAACCCCTCAT GGTAACTTTG TTAGCCCCGA TTTTGCCAGA TGAGCAAAGT 120
 GAGGTTTTTG AGGCTTAAAG TAACCTGCC AAGGTCACGT GGCTGGGAAG TAACTCTCCC 180
 30 AGTTCTGAGA TGTCCGAGCC TGAGCGTTT GTCAATGTAC ACCATCAACT CAGTGCTGCC 240
 AGTCATTCGA GCAGCCAGCT AGGTATGCA AGGTTTCTCC ACCTTAGCAC TGTGACATT 300
 35 TCGAGCCAGA TAATTCTCTG TGTGAGGAG CTGTCTATG CTTGTAGGA TATACAACAG 360
 CATCTGGCT TTACCCACCA GATGTTGAA CACCTCCCCA GTCGTGACAG CCAAAATGT 420
 CTATAGAGCT TGCCACGTAT ACCCAGGGT TCC 453
 40

45 (2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 459 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GTGACTGGCG CCTGCCCCG ACGATGTGG CCCCCTGTG TGTGTGTGT GTGTGTGTG 60
 55 CCGGCGGCGC CGGTCCCGAC CAGCAAGCC GCTCCCCACC CGGATGCTAA CAGCCAGGAA 120
 GCGCTTCAGA ACCTGCTCCA AAGAGTGGG GCTGGGGAG ACGGAGAGCT GCGGCGAGC 180
 60 TCACACCTGG CCCCAGGCTC TGGTGTATT GATGGGGCTG TGGTGGCCAC GCGACAGAA 240

AGCCGGGGAG GAAGACCTGC GGTTCCTGA GAGGCTCCA GGGCTGCAGG CCACGGCGAC 300
 AGGCTCCGGG GAACATGGGG CTTTCCTGT CCACTCCCA GGAGTGTGGG COTCAACGCA 360
 TTGGCAGGGG ACGGCCGTGT GCCCTCTYCA GACCCACCC CCAGATGCAT TTATTAGAAA 420
 TAATAAATTC TTTCTTAGCT AAAAAAAAAA AAAAAAAT 459

10

(2) INFORMATION FOR SEQ ID NO: 37:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 509 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

ATGAAATTTA CCACTCTCCT CTTCTTGCCA GTGTAGCAG GGGCCCTGGT CTATGCTGAA 60
 GATGCCTCCT CTGACTCGAC GGCTGCTGAT COTGCCAGG AAGCTGGGAC CTCTAAGCCT 120
 AATGAAGAGA TCTCAGGTCC AGCAGAAACA GCTTACCCCC CAGAGACAAC CACAACAGCC 180
 CAGGAGACTT CGGCGGCAGC AGTTGAGGG AGAGCCAAGG TCACCTCAAG CAGGCAGGAA 240
 CTAAACCCCC TGAAATCCAT AGTGSAGAAA AGTATCTTAC TAACAGAACA AGCCCTTGCA 300
 AAAGCAGGAA AAGGAATGCA CGSAGGCGTG CCAGGTGGAA AACAATTCAT CGAAAATGGA 360
 AGTGAATTG CACAAAATT ACTGAAGAAA TTCAGTCTAT TAAAACCATG GGCATGAGAA 420
 GCTGAAAAGA ATGGGATCAT TGSACTTAAA GCCTTAAATA CCCTGTAGC CCAGAGCTAT 480
 TAAAACGAAA GCATCCAAA AAAAAAAAAA 509

40

(2) INFORMATION FOR SEQ ID NO: 38:

45

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 598 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

ATGTTGGGCT GTGGGATCCC AGGCTGGGC CTGCTCCTGC TGCTGCAGG CTCGGCAGAC 60
 GGAAATGGAA TCCAGGATT CTTCTACCA TGGAGCTGTG AGGGTGACAT ATGGGACCGG 120
 GAGAGCTGTG GGGGCCAGGC GGCATCAT AGCCCCAACC TCTGCTGCG TCTCCGGTGC 180
 TGCTACCGCA ATGGGGTCTG CTACCACCAG CGTCCAGACG AAAACGTGCG GAGGAAGCAC 240

60

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100

ATGTGGGGGC TGGTCTGGAC GTCAGGGGC CTCCTCTCTC TGAGCTGCAG CATCTGCTTG
 TTCTGGTGGG CCAAGCGGCG GGAAGTGCTG CATATGCCCG GTTTCCTGGC GGTCCCGTGT
 GACATCTCCA AGTGGTCTCG GCTGTTCTCC AAGCACCGAG GAGCAGAGAA GATGCCGTCC
 ACGGGTAGGG TGCGAGTCTG CCTCTCCAAA GATGCCAGGG ATGTGAGGG AGGCACCGAG
 GGGGAAGGGA CGGAGGAGGG TGAGGAGACA GAGGCGAGGG AAGAGAGGA TTAGGGGAGT
 CCCCCGGGGA CTGCTCAATA CAGATACGGT GGACGGAAAA AAAAAAAAAA AAAAAAAAAA

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 454 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100

ATGGAGGCTG TTTTACAGT TTTTTTTTT GTTGTGTTT TGTTTTAAA GAATACAGAA
 GGAGCCAAGC TTTTTCGAC TTTGATCCA GCTGCAAGCT CAGGGCAGAG TCAAGGGCCT
 GGGTTGGAAA AACCTGACTC ACAAGAATGC ATAATTGACC CTTGCAGCTA CCAATAGCC
 CTTGGAGCTG GCACTGAACC AGGCTGCAAG ATTTGACTGC CTAAAAACA CAAGGCCCTC
 TAGGCCTGGC AGGGATGTCC CTGTGCCCAG CACTGGGGG TCGAAGACTG GTTCTAGCA
 CTACCGGTCA CGGCATGTC GTCTAGAAG GGTCCAGAAG ATTATTTTAC GTTGAGTCCA
 TTTTAAATGT TCTGATCACC TGACAGGGCA CCCCACCC CCACTCCA ATAAAAGCCG
 TGACGTTGG AAAAAAAAAA AAAAAAAAAA AAAA

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 425 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65
 70
 75
 80
 85
 90
 95
 100

GCTAAAGGCC ATTCCCTCCG CAGGGCATTT GCGTCGGGT GGGAGGGAA AACGCATCTT
 GTTAATTATT TTTAATCTA TTTATGTAC ATACCTGGG CAGGGGCTTG GGGAGGTGGA
 GGGGGRAGAA GGGTCCCCC TCTCTGCCCC TCCACTCCT TTTCTACGGC GATTGTCTG

5 TGTCTG3CCC CCACCCACTG MDCATCCCCC ATTGTGTCTT GGATGTGSTM CTATTTTTTA 240
 TCGSTCTCCT TTCCCTCCTT CCGCGTTCG GCCCCCGMCC CACCCCTGC TCCCACTACC 300
 CTTTGTCTCT TGTCTTTCTT TGGGTTCCTG TACAACCAA CTGTATACA CTGTGTACAC 360
 ACAACCAGYC WAAGGCAAAA CCGAACGGCA AACACTTTAA AAAAAAAAAA AAAAAACTGG 420
 10 GGGGT 425

15 (2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2471 base pairs
 (B) TYPE: nucleic acid
 20 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

25 GGCACGAGTA TGGCTTCCCG TGGACTCAGC CTCTTCCCCG ATTCCTGGCA CGAGGGGGCT 60
 TCGCGTCTGT GCTTCCTGTG GCTGACGTCA TCTGGAGGAG ATTTGCTTTC TTTTCTCCA 120
 AAAGGGGAGG AAATTGAAAC TGACTGGCCC ACGATGGGAA GAGGGGAAAG CCCAGGGGTA 180
 30 CAGGAGGCCT CTG3GTGAAG GCAGAGGCTA ACATGGGGTT CCGAGCGACC TTGGCCGTG 240
 GCCTGACCAT CTTGTGTCTG TGTGTCTGTA CTATCATCAT CTGCTTCACC TGCTCCTGCT 300
 35 GCTGCCTTTA CAAGACGTGC CGCCGACCAC GTCCG3TTGT CACCACCACC ACATCCACCA 360
 CTGTGGTGCA TGCCCCCTAT CTTCAGCCTC CAAGTGTGCC GGCAGCTAC CCTGGACCAA 420
 GCTACCAGGG CTACCACACC ATGCCGCCTC AGCCAGGGAT GGCAG3AGCA CCTACCCAA 480
 40 TGCAGTACCC ACCACCTTAC CCAGCCGAGC CCATGGGCCC ACG3CCTAC CACGAGACCC 540
 TGGCTGGAGA GCAGCCGCGC CTTACCCCGC CAGCCAGCCT CTTTACAACC CGCCCTACAT 600
 45 GGATGCCCCG AAG3CGGTCC TGTGAGCATT CCCTGGCCTC TGTG3T3CC ACTTGGTTAT 660
 GTGTGTGTG TGGT3ASTG GTGT3AGGC GCGGTTCTTT ACGCCCATG TGTGCTGTGT 720
 GTGTCCAGGC ACG3TTCCTT ACGCCCATG TGTGCTGTGT GTTCTT3CC TGTATATGTG 780
 50 GCTTCTCTG ATGT3ACAA GTT33GAAC AATCCTTGCC AGAGT333CT G3GACCAAGC 840
 TTTGTCTCTT TCCTCACTG AAATTATGCT TCCTAAAATC TCA33CAAA CTCAAAGAAT 900
 55 GGGGTGGTGG GGG3CACTCT GT3AGGTGGC CCCTGAGAGG T3333CCTC TCCAGGGCAC 960
 ATCTGGAGTT CTTCTTCAGC TTA33CTAGG GTGACCAAGT ACG3CTGTG ACACCAGGGT 1020
 60 GGCGCAGCTT TCT3TGTGAT G3AGATGIGT CCTGGTTTCG G3AG3GTACC AGCTGCTGCT 1080

TGAGGGCCATG GGTCCGCTCC CGGATTTGGG GTATTCGCTT GAGAGGCTAG GGACATGATG 1140
 CAGGGGAAAT TGGGATCTG GCGAAGTTGG ACTTNGATCC TTGCGGCAGA TGTCCCATTC 1200
 5 CTCCCTGGAA GTTGTCTATG CTGTTGCGGA TGAGGAGAC TCTTGATGCC AGAACACCTC 1260
 AGGCAGAGGC CTATCAGCT GTACCTCTCT GCTGSACTG TCCCTGTCC CCGCATCTCT 1320
 CCTGGGACCA GTTGGAGGGC CACATGCACA CACAGCTAG CTGCCCCAG GGAGCTCTGC 1380
 10 TGCCTCTCT GTCCCTGCCC TTCCACAGG TGAGGAGGTC TCTGTCCAC CAGCACACTC 1440
 AGTTCCTTC CTTCAGTGT TTTCATTTA TTTAGGCAA ACATTTTGCC TGTTTCTGT 1500
 15 TTCAAACATG AATGTTATA TGAGACTGAA ACCCTGCTT TGTGGAGGGA AATTGGCTCA 1560
 GAGATGGACA ACCGCGCAAT TGTGAGTCCC TGCTTCGGA CACCAGCCTC ATGGAATATG 1620
 CAACAACCTT TGTACCCAG TCCACGCTGT TCTGGCAGCA GGGACACCTG GGCCAAATGG 1680
 20 CCATCTGGAC CAATGTTAG GTTGGGGCC CTGATGGCA GCTGTGGCC AGACATGAAT 1740
 ACCTCGTCTT CTTCTCTCT CTATTAAGT TCCACCAGAG CTGTCTTAGC TCAAATCTGT 1800
 25 TGTGTTCTG AGTCTAGCT CTGTACACTT GTTATAATA AATGCAATCG TTTGGAAAA 1860
 AAAAAAAAAA AACTCTAGT GGGGGGCCC TACCCAATGG GTTCMARAT AGTAGARWAC 1920
 RAAAAAYMCA ANTGCACCA AAGAGGGGCC AGCGGANTTT TAAGAGGGCC CCTTTTGGG 1980
 30 GGNATCCANT TTAGTCGGG TTNTAAGG AATTTGCTG GGGGGGTTA GGGCCCGTT 2040
 KYTWCTTCCA ACCAAGGCTT YTYGTGTTA GGGCGGTTG GGGCCMATGG GCTGGGCTG 2100
 35 GTAAAGTGCT GGSIMAYTC MATTGAGTAG GGTGCTGCTG GCATTCCTGG CTGAGGCGG 2160
 ATGGTGTGCT AGGCTCTGTA GCTTGTCTCA GGTAGCTG GGGGCACACT TGGAGGCTGA 2220
 GGATAAGGGG CATECACCCA CAGTGTGGA TGTGTTGTT GAGACAACCG GACGTGCTCG 2280
 40 GCGGCACGTC TTSTAAGGC AGTAGCAGGA GCATGTGAAG CAGATGATGA TAGTGACGAC 2340
 AGACAGCACA AAGATGCTC AGTCAACGGC CAATGTGCT CCGAACCCCA TGTTAGCCTC 2400
 45 TGCCTTCACC CAGAGGCTC CTGTACCCCT GGGCTTCTC CTCTCCCAT CGTGGGCCAC 2460
 TCACTCGTGC C 2471

50

(2) INFORMATION FOR SEQ ID NO: 42:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2659 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (E) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

	GGCAGGAGCT TTTCTCTAGA GTCTGAAAGA TCTAGAAAG AAATAAAAT TAACTTACTT	60
5	AAGAGAATTA TGGATCTTTT ATTAATAAAA ATTAACCTGA TGATTTGAAC TAACAGTTAT	120
	GATAATTCTG GTATTTATAG CTTTTTTTAT TCCCTGCAG AAAACCATAG GCAAAATTCG	180
	AACATGCTTG GAATGCGAA GTGCAGCTTT ACAGTCCACA CAGTCTCAAG AAGAATTTAA	240
10	ACTGGAGGAC CTGAAGAAGC TAGAACCAGT CTTAAAGAAT ATTCTTACAT ATAATAAAGA	300
	ATTCCCATTT GATGTTGAGC CTGTCCCAT TAAAGAAGAT TTGGCACCTG GTGAAGAAGA	360
15	GAATTTGGA TTTGAAGAAG ATGAAGAAGA GGTGTGTGCT GGAGCAGGTC TCCTGATTCT	420
	TTCTTGCTAG AGTTCCCGGT ACTTTATTAC CAAGGTGACC ATCGGAACCA GGAATGACAT	480
	TACTCACTAT CAGAATTGAG AAAATTGGTT TGAAGATGC TGGGCAGTGC ATCGATCCCT	540
20	ATATTACAGT TAGTGTAAG GATCTGAATG GCATAGACTT AACTCCTGTG CAAGATACTT	600
	CTGTGCTTC AAGAAAAGAA GATACATATG TTCATTTTAA TGTGGACATT GAGCTCCAGA	660
25	AGCATGTTGA AAAATTAAAC AAGGTGAGC CTATCTTCTT TGAATTCAAA CACTACAAGT	720
	CTAAAAAAG GTTTACCAGC ACCAAGTGT TGTCTTTCAT GGAGATGGAT GAAATTAAAC	780
	CTGGGCCAAT TGTAAAGAA CTATACAAGA AACCCACTGA CTTTAAAGA AAGAAATTGT	840
30	AATTATTGAC CAAGAAACCA CTTTATCTTC ATCTACATCA AACTTTGCAC AAGGAATGAT	900
	CCTGACATGA TGAACCTGGA ACTTCTGTGA ATTTTACCAC TCAGTAGAAA CCATCATAGT	960
35	TCTGTGTAGC ATATTACCCC TTCAACAGGC AGGAAGCAAG CCGTACCCAG ACCAATAGGT	1020
	CGGACGAGT CAAATGCAAA GCTGTACCAC AGAATTCAGA GTCCAGCACA TCACACTGAC	1080
	GTATAGGACT CCTTGGGATA CAGGTTTATT GTAGATTTTG AAACATGTTT TTACTTTTCT	1140
40	ATTAATGTG CAATTAATAG TCTATTTTCT AATTACCAC TACTCCTACC CTGCTTCCTG	1200
	GAACAATACT GTTGTGGTA GGATGTGCTC ATCTTCAGAC TTAATACAGC AATAAGAATG	1260
45	TGCTAGAGTT TACACATCTG TTCACTTTTG CTCCAATATG CTCTTTTGAC TTAACGTCAA	1320
	GCTTTGGGTT GATGTGGTA GGTAGTCTC AAACGTCTT GAGAGGAATG GGACCAGTTT	1380
	TGCTGCTTAA GAAGTCTGT CTGGATGTTT ATAGGCAGCA CCTCTGAAGT GGCCTAAAT	1440
50	CACCTGATC TGATAGTTTT CCTGCTTAGA AAGTGTGCTT TGGCCAGATC AGTATCCAC	1500
	ATGGAAGTGT TCCCTAGGTT GTAGCTGTGA TTGTTCCAG ATGACCAGAT TGTTTTCTG	1560
55	AAAATGAGCA TATTTTAGT CATGTGATT AGGTGTTCTT CTACATCACA TTGTTACTCT	1620
	TTCTGATGAT GATTCTAGGG TTAACATTGG AACATCTCA AAATAATTAC AAAGTTTATG	1680
	ATGGGTTTAC AATGCTTCT AAACAATGTA ATCTAAAAAT AATTGAGTCA GATGCTAAC	1740
60	AGATACTGCA GGCATAACTG CTGTTTTTCT GACAACTGAT TGTGAAACCT TAAACCTGT	1800

	ATACCTCTTC TTACATGAG GATATGCAA AATCTGGAAA GATATCTAT TTTTCTTATA	1860
	TAGGTAGATA GATGAGCAT TTATTTCTTA TTATATATA CTGACATTA TCCATATGAA	1920
5	AATATGCAGG TATTAGCTT ACTATAATTT ACTTTTACT TAATGGGCA TAAATAAAC	1980
	TTTCATAGTA CACATAGGT GATATTTGA TACACAGAAC ATTTGGGTG GGTTTTCTGT	2040
10	GGTTAGATG TAAAGGACAC ATATTTCAT ATTCATATT TTAAATGAG AATGATGAG	2100
	GGGAATGCAG TGTATACC TGGCTATTT TTAACTAGT GTATCACCC TATCATACC	2160
	ATTCAGTATG TTGCTTTT AAAATAAGTA ACCACAATTA AGTTGTGTA GGCCTTGAC	2220
15	TTCAAGASAT CTAGTCTTA CTTCAGTTG TGTCTTAGGT CCATTCTGT TACTAGACG	2280
	ATTTAATAA AACTATGCG AGCTGGAAT GGAATCTCC AGCCAAATTT TAGTCTTGT	2340
20	CTTCCATCT TGATGGATT AATTCCAAAT TCTAAATGA TTCAGTCCAC AATAGCTCTA	2400
	GGGATGAAG AATTGCCCT ACTTTGCCA GTTCTAAGA CTGTGAGTT TCAATCCCT	2460
	AGACTGTAAG CTCTCAAGG AGCAAGAGG GCATTTTCTC CGTGCATGT AATTTTCTA	2520
25	AGTGTGTTG CAGCACTCTG TACCTGTGG AGTACTCAGT ACCTTTTGT TGATGTTGT	2580
	GACAAGACCT GAAAAAAT CCTTAAAA AAAACCCAT TAAAGTGTG CAAAACCGA	2640
30	AWAAAAAAA AAAAAAA	2659

35 (2) INFORMATION FOR SEQ ID NO: 43:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1635 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

45	CGAGGAGGTC ATGAACAAGG AGGCGGGAGA GGTGGACGTG GTGGTTATGA CCATGCTGGC	60
	CGAGGGGAG GAAGAGAAA TAAGCATCAA GAGGCTGGA CAGATGAGG GAGTGGTGA	120
	GGAGGTGCT ACCAGATGG TGGTATGTA GATTGAGGT TCCAGCCAGG TGCTATCAT	180
50	GTGCGCACCA GCACTGGTGG CTATCAAGGC GAGGTTATG GTGGTTCCA AACATCTTCT	240
	TCTATATACAG GAAGTGATA CAGATGCT GCTACCAGC AGGACAATAG ATACCAAGAT	300
55	GGCGGGCACC ATGCTGATCG TGGTGGTGGT GTGGTGGGC GAGTGGTGG TGGAGCCGA	360
	GGTGGTGGT CAGGDCAGG AGGAGGCTGG GAGGAAGAG GAGGCAGAA TTATCACCAA	420
60	GGGGTCAAT TTGAACAGCA TTTCAGCAT GAGGTTATC AGTATAATCA TTCTGGATTT	480

	GGACAGG3AA GACATTACAC TASTTSAGG3 TACCGAACCT TACATTTTGT TABAGCTCAA	540
	GTAATAGAAA CTTAGTTTCA GAATCCTGAA TTCAGCACCT ATTTTGAATT AATGTBAGAC	600
5	CACAGTGGC AGGCAGATTC CTGCTTG3DA TAAGCATTTG TAGGTCTTCA TTCAATTCTG	660
	TTAGATTTTT TTATGGACT TACATAATGC CGTTTATTTG AGAAACACAT AADATCTCTC	720
10	CTTCTATGA AAAATTTTTT AAAAG3T3BT TAAAATTGCC TTAAATG3CC CATTABACTA	780
	ATTCCACAGT CAGAACATGC AAACTTTTTT GAAGAAATTA CTGAATAAG TASTTTTCAI	840
	GTTTTCAATA TGCAGTTTGT AAAAT3AG3A TTCACCTAGA CTTTTTTAGA TTTACTACYA	900
15	GGAAACCTTC CYCATATGAA TAACCATTTA TATGTGTTTT GTTAAAGTA TTCCAATGCC	960
	TATTTTCCAA GCACAGTTCT G3CCCCG3BT TGACTTTTAT GCCACGT3TG CTTCATGATG	1020
20	GAACTTTATG GTCAGTTCTT ATTAATAGAG CTCTTGTGCA GATAGCACAT TCAGTAGCCT	1080
	TATTTTGTG ATGGAATACT GTATCATATG CTCAACTCTG AAAACCTTGA ACACG3CCAA	1140
	AATCCATAAA GATTATAAAA GCAAACTAAG TTGTGAAGCT ATAGTACATG TAGGCATTTA	1200
25	GTTAAGTATA GCAATTCAAA CTGACCTGCA TCCATCCAAA ACAAATTCCT CCTTCAACCT	1260
	TATTTTACT TGAAATTTGC TAGAAGAAAT AGCAAACCGA AATTTGTTTT ATGCATGAGT	1320
30	TAATACCACT G3TTCAGCAA ATACAAGTTA GTTTGCTTTA AGCAGGTAAC TTTTMTTGT	1380
	ATGGAAGAAA TGCCTACAA AGTTAAGACA GATTTTGTCT AAGTGCAGGA G3CCCTTTAT	1440
	TATTGTGCA GAAAACAAAA GCCTG3CTGA GTTGATGTTT TACATTCTCC CTTACTGAAA	1500
35	TCTACAT3AC ATGATGCTTC TTGCTG33TT TTTGTACATG TAAACATTGT CAAGCTGTGA	1560
	AAGAAAATGG CTGAGGTGT GCTTT3T3TG AAAG3TGAGC ACTGAAAGTA TCT3TTAAGT	1620
40	TCTCCNGAAA AAAAA	1635

45 (2) INFORMATION FOR SEQ ID NO: 44:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 780 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

55	AACATGTCAT TGTCTTTTAG TTTCATTATT TTCTACTCC TTGTATGTCA AGAAATTACA	60
	TTTGTATGT CTATGGAGA TGCTTTAAT TGCTTCAGTG AGTGCTTTTC TAATCTGCAG	120
	ACCAATTACA TTTCCTGPTT G3AGCAT3CT GT3TGCAAAC ATTCAGTAAT TTG3ATAT	180
60	CAATTATTTG TTAGGGCTCT TCCTATTTCC AAATGTGCTG AATTGTCTAT T3ATG3GATT	240

TTCATATCTT TTGATGAGAA CTGGAATGT AGCTGGGTGG CAGCTATCTA GGTTCCTACG 30
 TATGAGTAG ACTTCTCTT GGTATAGTA AGCTCAGAC AGTTTACT TTTATCTACT 36
 5 TTACTTGTGG AATAAACA GTGATTTGT TTTAAAGAA TAAGTAGCT TTCTGTAGAG 42
 AAGGATTC TACCTCTAAA AGCTGCTTG AAGCTCAGA ACTGCACTT TTCTAGGTG 48
 10 ATTTTAAAT TTGATATTA GGGAGAGTC AGCATTTGCT GACACAGATT CTACATACT 54
 AATGTATGAT AGCAATGCA AACATATAT AATGTGTGT ATTTGTGCA TACACAGGT 60
 AGACAACTA GACTCTGGCA GCAGATCTCC AGACACCAA GTTATGTTT TCATAGTGA 66
 15 TTGAGTAG TTATCTCTT GGTATAGTA GTTATGTCC TGGAGAACT CATTACTGAA 72
 AAGCATTTAA CTAAAAAA AAAAAAAAA AACTGAAA AGTACTGAA TACAGATAG 78
 20

(2) INFORMATION FOR SEQ ID NO: 45:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2378 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

GCGAAGCAGC TGAAGCCGCG GCGGCGAGA ATCCAGCTG GCTTCGCGG CCAAGGTCAC 60
 35 CCACAGCAAG TTCCCGCGG CGGATAGAG CCGCCCCCTG GAAAGAGGC TCGGCTCAA 120
 GACCTTCAGC TCCAGAGCG ATATCAGCT GGTGTAAAC GAGTACCA AATGCAGGAG 180
 40 AGCGGCTTCT ACTGAGGCG ATGAGCGCG GAGAGAGA AGTTCTCT CAGTCCCGAG 240
 CCGGCGGCA CCTTCTGAT CGGTACAGC TGGGACCGAG CGCACTTCT TCACGCTCAG 300
 CGTCAAGACC CAGTCTGGA CCAAAACCT GCGCATCCAG TGTAGGAGG GAGCTTCTC 360
 45 TGTGAGAGC GATCCCCGA GAGAGAGGC GTGSCCGC TTAGACTAG TGTCAAGCT 420
 GTTGACCAC TATATGCGT CCGTAGAGC CCGTCTCTT CCGTGGGAC CTACTGAAC 480
 CTCTCCGAG GTGCGGAGC AGCTCTGTG CCAGCCACTC CTGGAATG CCGCAGAG 540
 50 AGCTATTAC ATATCTCTG GGGAGAGAA GATCCCCCTG GTGTAGAG GGTCTCTCT 600
 CTCAACGTG GCACTCTCT AGCTCTCTG TCGAAGAGC GTTAATGCG AATGGACTC 660
 55 CTATGAGAAA GTACCCAGC TGGGCGGCG CATTCGGAG TTAGAGAGC AGTACGATG 720
 CCGCTTTAA GGGTAAAGG GCGAAAGG CATGGGTGG GAGAGGGAC GCAGGCCCT 780
 CTCTCCGTG GCATAGGCA CAGGACAAAG AACCAACCA GAGAGAGTC CTATAGCTCT 840
 60

GGGGGGAAAG AGGGGGGACA GGGGGCTCCC TCTGCCCTCT CCTGCGAGAA TGTGGCAGGC 900
 GGACCTGGAA TGTGTTGGAG GBAAGGGGA GTACCACCTG AGTCTCCAGC TTCTCCGGAG 960
 5 GASCCAGCTG TCTGCTGGG AGGATAGCAA CCACAAGTGG ATTCTCCTTC AATTCTCAG 1020
 CTCCCCCTCT GCTTCCAAAC AGGGGACACT TCGGGAATGC TGAACATAATG AGAACTGCCA 1080
 10 GGGAACTCTTC AAACCTTCCA AGGGAAGTGG TTGCTCTTT GATTGCTTT AAACCTGAGC 1140
 TGGTTGTGGA GCTTGGGAAA GGTGGAAGAG AGAGAGGTCC TGAGGGCCCC AGGGCTGCGG 1200
 GCTGGCGAAG GAAATGCTCA CACCCCCCGC CCACCCAGG CGAGSATCCT GGTGACATGC 1260
 15 TCCTCTCCCT GCTCCGGGG AGAAGGGCTT GGGGTGACCT GAAAGGGAAC CATCTCGGTG 1320
 CCCCACATCC TCTCTCCGG GACAGTCACC GAAAACACAG GTTCCAAAGT CTACCTGGTG 1380
 CCTGAGAGCC CAGGGCCCTT CTTCCGTTTT AAGGGGGAAG CAACATTTGG CACGAGATGG 1440
 20 GCTGGTCAGC TGCTCTCTT TTCTACTCA TACTATACCT TCCTGTACCT GGTGGATGG 1500
 AGCGGGAGGA TGGAGAGAGC GACATCTTT CACCTCAGGC TCCTGCTAGA GAATACAGGG 1560
 25 GATTCTACTC TGTGCTCTT GACTATGTCT GGCTAAGAGA TCTGCTTAA ATGCTCCCTG 1620
 TCCCATGGAG AAGGACCCAG CATAGGAAAG CCACATACTC AGCCTGGATG GGTGGAGAGG 1680
 CTGAGGGACT CACTGGAGGG CACCAAGCCA GCCACAGCC AGGGAAGTGG GGAGGGGGGC 1740
 30 GGAAACCCAT GCTCCAGC TGAGCACTGG GAATGTCAGC CCAGTAAGTA TTGGCCAGTC 1800
 AGGCGCCTCG TGGTCAGAGC AGAGCCACCA GGTCCCACTG CCCCAGCCC TGCACAGCCC 1860
 35 TCCCTCCTGC CTGGGTGGG GAGGCTGGAG GTCATGGAG AGGCTGGACT GCTGCCACCC 1920
 CGGGTGCTCC CGCTCTGCA TAGCACTGAT CAGTGACAAT TTACAGGAAT GTAGCAGCGA 1980
 40 TGGAAATTACC TGAACAGTT TTTTGTTTT GTTTTGTGT TTGTTTTGT GGGGGGGGGC 2040
 AACTAAACAA ACACAAAGTA TTCTGTGTCA GGTATTGGGC TGGACAGGGC AGTTGTGTGT 2100
 TGGGGTGGTT TTTTCTCTA TTTTCTTGT TGTCTTGT TTTTAATAA TGTTTACAAT 2160
 45 CTGCCTCAAT CACTCTGTCT TTTATAAAGA TTCCACTCCA GTCCTCTCTC CTCCCCCTA 2220
 CTCAGGCCCT TGAGGCTATT AGGAGATGCT TGAAGAACTC AACAAAATCC CAATCCAAGT 2280
 50 CAACTTTGC ACATATTAT ATTTATATTC AGAAAAGAAA CATTCAGTA ATTTATAATA 2340
 AAGAGCACTA TTTTTAATG AAAAAAAAAA AAAAAAAAAA 2378

55

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1772 base pairs

(B) TYPE: nucleic acid

60

(D) STRANDEDNESS: double
(E) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46

5	TCGACCCACG CTTCCCGGGAAT GATCCACAGC CGGGTCCCAA GCTCTGTGCT GAGCCTGAGC	60
	CTGAGCTTAA GTCGATCGG GAGCGCTTCG CCGGGGCTCC GGGCTGTGCG ACCTGCTGGG	120
10	CCCCAGCGAT GATGACGCTG TCGGAGGGCC TTCTTCGGCT TGGTTCCTTG CTCAGCCTGT	180
	CGTCTCTGAC GTTTCGCTG CTGCTCTGCG CCGACTGTCA GAGCTCCTCA AGAATTTTCA	240
15	GGATGTCAGA TTTAAAGTAA TCTGCTCTCC CTATAAAGAA AAATCTCTGG CATATTTTATA	300
	ATAAGAACAT AATCTAGAAA GATTCTGATT GCTTCATCTT TCTGAGAGCC ATGCTCTGTG	360
	GGGGTCTTGA TGTAGAAATA TACTCTCTAC GTCTGGAATG CAATATGAA GAAAGAAGCT	420
20	CTGTACAAAT CAAGCTTACG ATTATAATTT AATCTCTCAT TTCTGCTCTT CTACTTCTGT	480
	ACATGCTATA TATTACTCTG GTTGAGTCCA TACTGAGAG GGGCTCTCTT GGACATGCA	540
	AGTTGATACA GATGATGAT GATATTGGGG ATCACCAGCC TTCTGCAAAAT GCACACGATG	600
25	TGCTAGCCCG CTCTCGGAGT CGAGCCACCG TGCTGAACAA GGTAGAATAT GGCACAGCAG	660
	CGCTGSAAGC TTTAAGTCCA AGAGCAGCGA AAAGTCTGTC TTCTACCGGC ATGTTGTCTT	720
30	CAGCTAATTG GGTAAATGAA TTCAAGCTGA CTAGAAAGAA ACAGTCAGAC AACTGGAAAG	780
	GAACTGACTG GCTTTTCTG GCTTCTATTT TAATACCTTG TTGATTTTAC CACTGTTG	840
	TGGAGGATTC AAAATGGA GAAAAAATT GCTTGATTTT TTTTCTCTGT TAACGTAATA	900
35	ATAGAGACAT TTTTAAAGC ACACGCTCA AAGTCAGCCA ATAAGTCTTT TCTATTTG	960
	GACTTTTACT TATATAAATA AATCTGCTG TAAATAAAT TAAAAATCC TTTACCTGSA	1020
40	ACAAGCACTC TCTTTTTCAC CACATAGTTT TAACTTGACT TTCTAAGATA ATTTTCAGGG	1080
	CTTTCTGTG TCTTTTCTTT TCTTTCTTGG TTTTCTGCG AGAGGCGAGG GATGCTCTGG	1140
	AAGTGCTTAA CACTTTCTTT CAAGTCACTT TACTAAGCAA ACTTTTGTAA ATAGACCTTA	1200
45	CCTTCTATTT TCGAGTTTCA TTTATATTTT CCACTGTAGC CAGCTCATC AAAGAGCTGA	1260
	CTTACTCATT TCACTTTTGG ACTGATGTA TATCTGCTT ATCTGCTGTG TCTGCACTT	1320
50	ATGCTAAAGC GATTTTAAAT TCCCTGCTGG CTTTTCACAA AAAGCAGATT TCTTCTATCT	1380
	ACTGTGATGT CTGATGCAAT GATCTCTAGA ACAAAGCTGG CATTTGCTAG TTTACTCTAA	1440
	AGACTAAACA TATCTTCTGT GTCTCTGCTG TTAATCATCT TCTACTACCT TTAAGSACAA	1500
55	ATCCTAAGGA CTGAGCACT TCGAATAAAG AAATTTTATT TTAAACCCAA GCCTCCCTGG	1560
	ATTGATAATA TATACACATT TCTCAGCAT TCCGCTCTG CTGAGAGGCA GTGTCTGAG	1620
60	CTCCAATGTG TGCAGCTTTG AACTAGGCTT GGGGTTGTGG GTGCTCTTTC TGAAAGGTCT	1680

AACCATTAAT GGATAACTGG CTTTTTTTTT TCCTCTTTGG AATGTAACAA TAAAAATAAT 1740
 TTTTGAAACA TCAAAAAAAA AAAAAAAA AA 1771

5

(2) INFORMATION FOR SEQ ID NO: 47:

10

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1107 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

CGGGCGAGAA GGGCAGACGG GACATGCAGC CTCTCCGCC TGAGCCCCGG AAGTGATGTG 60
 GCTGCGGCAT CGCGGCCTCG CTATGCTGCG CATTTCAT TTTGAGAGTC TATTGACTGT 120
 AATCTTGCTG CTTATATGTA CTTGTCCTTA TATTCGATCC TTGGCACCCA GCCTCCTGGA 180
 CAGAAATAAA ACTGGATTGT TGGGTATATT TTGGAAGTGT GCCAGAATTG GTGAACGSA 240
 GAGTCTTAT GTTCAGTAT GCTGTATAGT AATGSCCTTC AGCATCCTCT TCATACAGTA 300
 GCTGGGSAAA ATGCCAGAAT GTAGTGGCCA TCAGATTGA TTGTGAACAA GGACTGACTG 360
 CAGAAATAA TGAAGAGGAT GTTTAACTCT TTTATCTCCG AACATTGAAT GAGATAAAT 420
 TCCAGATGCT GTTCTCTATT TTAATCTTAT TGGACCAATG TTCTGTATAA ACAATTAAGA 480
 TGTAACCATT TAATAGTCTG TAACAATCAA CTTGAGTACT GTCAGTACAA TATTACATT 540
 TGCAAAATTT ATTCTGTTGT ATCAGATACA AAATTTTAGT GAGGTATCTC TAAGGCACAT 600
 AGTAGAAAAC AAAATTGGTT AATTACTCAA GTTCCTTTCA CTGTGATTG GAAATGATTT 660
 AATCTTTATA GAATGAGAAC CTTTTTGGG CTAGCTTTT TATTAAAATG GCTCAATTTG 720
 TGTGATAAG GATTGCATTA ATATTTAATA GTGCTTGCTT TTCTCTGGG CACACCATTT 780
 TGATCATTA CCAGAGTACC TCTACTCTTA GAAAACCTA GTTATGACA AGTATTTAAA 840
 ATATTTAAAA CAAGCTTATG CAGTCTTAA GAGGAAGGT AAATGAGATG TAACTTAAAA 900
 ATAGTATTTG GAAAATGTTG ATAGTTAACA TTAGTGGATT TAGACTAGCC AAATGACATA 960
 GTAGGCTCTG AACATCTTG TCAATATAT GIATTTTGTG CATGAATTTT TGCTGGAAAG 1020
 CTGTCTTTCT CTGAAAAACA CAACGTTCTT AATATGAAAA GAACAATTAT AAAATAAAA 1080
 AAAAATTTAA AAAAAACTGG GCGGGGG 1107

60

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 805 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

10 TGCAGAGAG AGGAGTTGC TTTTGGAAA CTACTACCGA TTGGCTGAGC ATCTCTCCAA 60
TGCAGTTGGT GAGGTTAGGG TTTTSATGGA TGATTCACAA AGTATTATTT GATTAATCT 120
GGACAGGCGC CGAAACGTGA TGATGAGGTI GAATCTACAG CTGACCATGG GAAGCTTCTT 180
15 TTTTGGGTTT TTGGACTAA TGGGASTTGT TTTTGGAAAT AATTIGGAAT CTTCCTCTGA 240
AGAGGAGCAT AGAATTTTCT GGTGATWAC AGGAATTATG TTCATGGGAA GTGCTCTCAT 300
20 CTGAGGGBCT CTCTTTTCAT TCTTGGAGCG ACAGTTAGAA GCTTCATTCG CCGCTATGGT 360
ATGAAAGGATA TGGTTCACGG CGGATTTGTT GAAGGTTTAT GATCATGGGC GTTAAAGTCA 420
GAGGCTCTGG GATTAAGTTG TCACAGGCGC TATGGCCCTT GGGAGTTGCT TTCTCAAACI 480
25 TCTTTCAGTT TCGCTATCTG TCACTTAAAT CGGATTTACC TGGTTCATAG GTTTATGGGA 540
AGAATTAAAC AATATGTGTA AAGCACTTAC TAGCACACTG CCAACACAAA TAAGTTAGAA 600
30 ATATATTTTG TGTAGAATTC TGACAACATA CATTTAAACA GATTTTAAAT ATTCTGGTAT 660
AAGGTTTCTC ATAACCAAT GGAATGTAT GAAACATTTA TAATGTTCTT AAAAGATAGT 720
AAATTCACCT CCACTTTCTT TGTACTTGAA GATGGCACCA CTGGAATAAA TACTTAAGAC 780
35 ACTGAAAAAA AAAAAAAAAA AACTT 805

40

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1408 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

50 TCATTATTTA TTCATGTGGC TGAAGAGTA TATTAATTAT GTTAGATTT TTGAAAAAG 60
TCTGAACAAA AAAAGGACCT ATACAGTGCT CAACTATAT TTTTAAAAAT ACTATTTTAT 120
55 TTTTACTCAC ATATGAAAAA AATGGGTGTA CTATCATGTT TACATACATA CTAACATTGG 180
AAACAGAATA ACGAATTGTA TTTAAATTTT ATGAAGAACA CACAAACATT AAAACACTGA 240
TTGGTACAG AAGGAGAST TTGAGGAAAA AACATTAGCT ATAATTTTCA TTTTCAATTA 300
60

AGAGCAGCAC CCTCTGAGAA TAATCAAACT GATTASTAAT ATTCACTAT ACTGCAAAAT 360
 AATATGTACA AAGGAAAATT ATGATTCTA CTGATTTTAT TACTTTIACC AAGCCATTTT 420
 5 ATGTTCTCTCA CTCAATGCAA AGAAATAAAA CATAACTCTGA AGAAAAATAT GTCTTTATTA 480
 TTATTTACAA TAAAAAGTTG GCTTTACTCT GCAAGCCTGG GCATATTSTA CAATTGGCAG 540
 10 CACTTAACGG CTCAAGTGGT TCAMGTACC AGTTTGATTC TGATCCACTG AATAGAATCT 600
 CTCATTCATA TCTGGTGACC AGACTAACTC CATGGGAGCT GTGATAGACT GAACCATTTT 660
 TGTGGTATCC CTAGATCTCA CTAATAAGA AAGATCCTAC ACCAGAAAAT ATAGCAACTG 720
 15 ATCTATCTAT AAATTACATC TATATGCTAG CTCTTTAGTA TAAGTTGGAA AAAGGGGCCC 780
 TTTCTTGAGC ACATGGATAA AAGTATTTT GTAGTCTAAA GATTGCTGGA TTGATATTGT 840
 GTTCTTATAA TGAAGATAAG CTACACACTG AAACCACTGT CAGATTAAGA AACTTCCACA 900
 20 ACTTGTCTCA GTTCTTCAA CAATGGAACA AGTTCTTTT CTAGGCTGAC AATTAGTCTT 960
 GTATTGGCAC TGCTGCTGGC TATGAAACTC ACCACCAAAG GTAAACGATT AAATTGAACC 1020
 25 ACCTGCTAGG TGTATAGTA ACAGATGATA CTTTATTTT TGGAAAGTCC AAGTTTGCTT 1080
 CCTTGCTCTG TTGCAAGGGC AAAAGTGBAT AAGAAACCAG GTCGCAAAGC ATGCTCTGGA 1140
 GCATTGTCTAT TGGCCACTTT AATAACAGGT ACTCCATCTC TATCTGACAC AACAATGGCA 1200
 30 TGBAGCCCTT CAACACTTGG TAACTTTTTA TACAAGAATC GCTTTAGGTC ATCCGCCATG 1260
 ATGAACCCCC TTCTCTGCA GGATCAATCT CCACGCCTGG GGTTCCTGGG CTGCCTGGTT 1320
 35 CTCTCCGCTG TCACTTCAGG GACAGCTTTA AAGACAGGTT CCTCCTCAAG CCACGCTCAC 1380
 ATGATTCATG ACCTGCTCTG CGCTCCAG 1408

40

(2) INFORMATION FOR SEQ ID NO: 50:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1813 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

CATGGTGGGG CACGAGATGG CCTCTRACTC TTCWAACACT TCACTGCCAT TCTCAAACAT 60
 GGGAAATCCA ATGAACACCA CACAGTIAGG GAAATCACTT TTTCAGTGGC AGGTGAGCA 120
 55 GGAAGAAAGC AAATTGGCAA ATATTTTCCA AGACCACTTT CTTTCAAAGG ATGCAGATGG 180
 TGACACGTTT CTTCATATTG CTCTTCTCCA AGGAGAGAAG GCACTTTCTT ATGTTCTTGG 240
 60 AAGAAAGATG AATGCACTTC ACATGCTTGA TATTAAAGAG CACAATCGAC AGAGTGCCTT 300

10AGGTGGCA GTGGTGGCA AGCAGCATCT CATGTGGAG GATCTGTGA ACATGGGGG 360
 AATGTTGAC ACCACAGACT GTTGGGAG AACACCTTG CATGTGTGTG CTGAGAGGG 420
 5 CCAATTCAG GTGTTCAGG CATTTCAGAA GGGAGCASTG GGAAGTAATG AGTTTGTGA 480
 TTTTBAAGCA ACTAACTATG ATGGGCTGAC TCCCTTCAG TTGCACTCA TAGCCACAA 540
 10 TGTCTGTGTC CATGAACCTT AGAGAAATCA ACAGCCTCAT TCACCTAAG TCCAGSAGCT 600
 TTTACTBAAG AATAAGAGTG TGTGTATAC CATTAAGTGC CTAATTAAA TGGGAGCAGC 660
 GGTGAAGCG AAGGATCGCA AAGTGGGTC CACAGCCCTG CATTGTGAG CTGAAGAAGT 720
 15 AATCTGGAA CTCATTGCGG TCTTTTGA GTGTCCAGT TGGCTGTCTT TTGTGAATGC 780
 AATGCTTAC AATGGCAACA CTGCGCTCCA TGTGTCTGCC AGTTTGCACT ATCGTTGAC 840
 20 ACAATTAGAT GCTGTCCGCG TGTGATGAG GAAGGAGCA GAACCAAGTA CTCGAACTT 900
 GGABAAGGAA CAGCCAGTGC ATTTGCTCC GATNGCCCT CTGGGAAAC AGATCCGACG 960
 TATCTGAAG GGAAAGTCCA TTCACTAGAG AGCTCCACCG TATTAGCTCC ATTAGCTTG 1020
 25 AGCCGTGCTA GCAACACTCA CTGTAGTTA GGCAGTCTG ATGTATCTGT ACATAGACCA 1080
 TTTGCTTAT ATGGCAAT GTAACTTGT TCTATGAAC AAACATATT AGTTCACTAT 1140
 30 TATATAGTGG GTTATATTGA AAGAAAGAA AAAAAATATC TAATTWCTCT TGGCAGATT 1200
 GCATATTCA TACCCAGGTA TCTGATCTA GACATCTGAA TTTGATCTCA ATGGTAACAT 1260
 TGCCTTCAAT TAACAGTAGC TTTTGAAGT GAAAGGACTT TGATTTGTGG CACAAAACAT 1320
 35 TATTAATATA GCTATTGACA GTTTAAAGC AGGTAAATTG TAAATGTTTC TTAAAGAAAA 1380
 AGCATGTGAA AGGAAAAAG TAAATACAGC ATTGAGGCTT CATTTGGCTT TAGTCCCTGG 1440
 40 GACTTACTGG CGTTGGACAG GCTTCAGTCA TTGGACTAGA TGAAGGTGT CCATGGTTAG 1500
 AATTGATCT TTGCAAACTG TATATACTG TATTTTGT CTTTAAAAAT ATTGTACATA 1560
 CTGTGTGTT AACATGGTCA TATTGAAAT CTATAAGTCC ATAAAAAGA AAGGAACAAG 1620
 45 TGAATTGTG CTATTAAAA AATTTTACA ATTCTTACTA AGGAGTTTTT ATTGTGTAAT 1680
 CATTAAGTCT TTTAGATAA AGCAGATGGG GAGTTACGGA GTTGTCTCTT TACTGGGTGA 1740
 50 AAGATATATT CGAATTGTA AGATGCTTTT YCTATGCAT TGAATTTATA CATTATTGT 1800
 AGGGAATTGC ATG 1812

55

(2) INFORMATION FOR SEQ ID NO: 51:

(1) SEQUENCE CHARACTERISTICS:

60

(A) LENGTH: 2070 base pairs

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

	CCACGGGTCC GGAAGAGGCG GGAAGTTCG CTGCGCGCTG GGTGCTGCT GGTCTCTGGA	60
	GGTGGGGGCG GGAAGGAGG GGGGGGCTG CCGGGGACT GGGATTCCT GGTTCCTCT	120
10	CCACCCGACG CGGCTTGAC CATGAGGCG ASATGCTGGG CAGTGGTGT GTTGCTGCG	180
	TTCTCTCTCC TAGGGGAGG TGGGAGACT CCGAAGCC CTGCGAGTC ATGAGCCAG	240
15	CTATGCTTCT TCGATTCTT GGTGAATGCT GTTGGCTATG CAGCTTTAT GTTACCAGGC	300
	TACCTCTCTG TCGACTACTT CAGGGGGAAG AACTACCTGG AGACGGTAG GGGCTCTGC	360
	TTTCTCTCTG TGAAAGCTTG TGTGTTGGG AATGAGCCCA AGGCTCTGA TGAGTTCTCC	420
20	CTGGGGCTCT GAAGAGAGG GGAAGAGCC ACCCGATGT GGAAGGCTT GAAGCTGCT	480
	TTCTGTGACA CAGGGCTCCA GGTCTCTTAT CTGACTTGGG GTGCTCTCA GGAAGAGTG	540
25	ATGACCCGCA GGTATGCGGC CACAGCCACA TCACCGGCTG AGCGCTTTAC GGAATCGCAG	600
	TTCTCTGTGC TAAGGAAGCG AGTCTCTGCA CTGATTGTGG CTGGCTCTCT CTGTCTCTCT	660
	TGTAAGCAGT CCGGGCATGG GGAAGCATG TACCGTACT CTTTCTCTCA GGTCTCTCAA	720
30	TGTGCTTAGC AGCTGCTGCG AATAGGAAG TCTTAAGTTC GTGAGCTTCC CACCCAGGT	780
	GCTGGCCAGG GCTCTAAGG TGATCTCTGT CATGCTGATG GGAAGCTTG TGTCTCGCG	840
35	CAGTAACGAA CACTGGGAGT ACCGACAGC CACCTCATC TGCATGCGG TCAGCATGTT	900
	TCTGCTATCC AGCGAGCAG AGGCGCGTAG CTCCCGAGCC ACCACCTCT CAGGCTCAT	960
	CTTACTGACA GTTATATTC CTTTGAACA GCTTCACTC AACTGCGAG GATGCCCTGT	1020
40	TTGCTTATAA GATGTCATCG GTGAGATGA TGTGCGGGG TCAATTTCTT CTCCTGCCTC	1080
	TTACAGTGG GCTCACTGCT AGAAACAGGG GCGCTACTG GAGGGAACCC GTTCTATGG	1140
45	GCGACACAGT GAGTTGCTG CCGATGCTT GGTACTCTCC ATGCTCTCG CATGTGGCCA	1200
	GCTCTTCATC TTTTACACA TGGGCGAGT TGGGCTGCG GTCTTACCA TATCATGAC	1260
	CCTCCGCCAG GCTTTGACA TCTTTCTTC CTGCTTCTC TATGCTACA CTGCTACTGT	1320
50	GGTGGGAGGG CTGGGGTGG CTGCTGCTT TGTGCTCTC CTGCTCAGG TGTACGCGG	1380
	GGGCGCTCTA AAGCAACGGG GAAAGAAAGG TGTGCTCTT GAGTCTCTG TGCAGAGGT	1440
55	TTGAGGCTGG AAGGCTCTG AGGCTGAAG TCAAAATAGG CCGGCTCAG ATCCCTTCT	1500
	GCTGTAAGCT CTGAGGAGT TGGTGAAG GCAAAATGC AGGCTTTTC TCACTATCAC	1560
60	AGACCAGCTC TGAGGAGGG GATGAGGAG CCGAGAGGC AGGCTCTCT TTTGCTTAA	1620

GTGACCCATC TTCCAGTAAG CATTCTATTC TGAGTCTCGG GCGTACACAG TCCCTAGTGA 1680
 GGGCTTTTGG GGAATTTCAG GTTAAGAGAG CATATCTAGG TTCCAGCTTT ACTCTTCCCA 1740
 5 CAAGTTCCCT TAATCTCTGT CCTAGCTGCG CCGTCCGACG TTCCAGACTC ACTCCCTCT 1800
 GCAATATCCG GCATCTCTTA CCGTGTCTAG AAAAGACAAA GGGCTCTAGG CTCGAATGCT 1860
 GCTTTCCGAG GAGCTGAAG ATGGTGTCTT GCTGAGGAAA GAGGATGAG AGCCCTGCG 1920
 10 AGCACCACCA CCGCTATATC TCTGTGATCC CTAGGCTCTG TTCCATGAGC CTGTTGCAGG 1980
 TTTTGTATCT TTAGAAATCT AACTTTTTCG TCTTATAATT TTATTTTATT AATTTAAATT 2040
 15 ACTGCAAAAA AAAAAAAAAA AAAAAAAAAA 2070

20 (2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1416 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

30 CCTCACTAA AGGGAACAAA AGCTGGAGCT CCACCGCGGT GCGGCGCGCT CTAGAACTAG 60
 TGGATCCCCC GGGCTGAGG AATTGGGCAC ACCATCCGGG GTCCGAGCG GCGGCTGCT 120
 GAGCTGCCTT GAGCTGAGT GGTGGGATC CAGAGCCATG TGGAGCTGC TACTACTGGG 180
 35 CCTGATTGGG GGGCTGACTC TCTTACTGCT GCTGAGCTG CTGGCTTTG CCGGGTACTC 240
 AGGGCTACTG GGTGGGCTG AAGTGAAGTC TGGGTACCTT CCGATCCGCA ACCTCACTGT 300
 40 GGCTTACAAG TTCCAGATCG GGTCTATGG TGAGACTGGG CGGCTTTTCA CTGAGAGCTG 360
 CAGCATCTCT CCGAAGCTTC GTCCATCGC TGTCTACTAT GACAACCCCC ACATGGTGGC 420
 CCGTGAATG TCCGATCTG CCGTGGGCAG CATCTGAGT GAAGGTGAGG AATCGCCCTC 480
 45 CCGTGAAGTC ATGAGCTCTT ACCAGAAATT TGGCTTCAAG GTGTCTCTCT TCCCGGAACG 540
 CAGCCATCTT GAGGACCTCA CTTTCTGCT AAGACCAACA TTCTGTCCCA TCTGGCTGGG 600
 50 CTACCCGGGG TCTCATCTCT GCGTGGACA CTTACTTCAA GAGGTGGAAG CTGTGTGCTT 660
 ATCTCGGCT GGGATCTAC CAGGAAGACC AGAATTCATT TCACTGCCCC ACTGGCACGG 720
 CCAGGAGAC TTTATGTTCT CTGAGATGAA GAGAGAGAG TGGAAATGGC GGGGGCTTCT 780
 55 GGAGGCGATT GACACCCAGG TGGATGCTAC AGGAGCTGAC ACAATGAGTG ACACGAGTTC 840
 TGTAACTTG GAAGTGACTT CTGGCAGCGG GGAGACTTCA GTTGGCACAC TGTCACTTGG 900
 60 GCGGAGCAGC CGTGGCTGGG ATGACGCTGA CACCCGACGC GAGCACAGCT AACAGCGAGT 960

CAGGTGCCAG CGGTCCTCT TTTGAGGAGC TGGACTTTGG AAGGCGAGGG GCCCTTAAGG 1020
 5 GGAGTACGG CTGGACCTTG GGACTTGAGC CCTTGGGGGA CTACCAAGTG GCTCTGGGAG 1080
 CCCACTGCCC CTGAGAAGGG CAAGGASTAA CCCATGBCCT GCACCCCTCT GCAGTGCAGT 1140
 TGCTGAGGAA CTGAGCAGAC TCTCCAGCAG ACTCTCCAGC CCTCTTCTCT CTTCCTCTG 1200
 10 GGGAGGAGGG CTTCCTGAGG GACCTGACTT CCCCTGCTCC AGGCTCTCTG CTAAGCCTTT 1260
 TCCTCACTGC CCTTLAGGCT CCCAGGGCCA GAGGAGCCAG GGACTATTTT CTGCACCAG 1320
 15 CCCAGGGCT GCGGCCCTG TTGTGTCTTT TTTTCAGACT CACAGTGGAG CTTCAGGAC 1380
 CCAGAATAAA GGCATGATT TACTTGTAA AAAAAAAAAA AAAAAA 1420

20

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

25 (A) LENGTH: 1720 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

30

GGCACCAGTG CGGCCCCAGC CTCTCCTCAC GCTCGCGCAG TCTCCGCCGC AGTCTCAGCT 60

GCAGCTGCAG GACTGAGCCG TGCACCCGGA GGAGACCCCC GAGGAGGCG ACAAACTTCG 120

35

CAGTGCCGCG ACGCAACCCC AGCCTGGGT AGCCTGCAGC ATGCCCCAGC TGTTCCTGCT 180

CCTGCTGGCA GCTCTGCTCC TGGCCCAGGC TCCTGCAGCT TTAGCAGATG TTCTGGAAG 240

40

AGACAGCTCA GAGGACCGCG CTTTTCGCGT GCGCATCGCG GCGACGCGC CACTGCAGG 300

CGTGCTCGGC GCGGCCCTCA CCATCCCTTG CCACGTCCAC TACCTGCGGC CACCGCCGAG 360

CCGCCGGGCT GTGTGAGCT CTCCGCGGCT CAAGTGGACT TTCTGTCCC GGGGCCGGGA 420

45

GGCAGAACTG CTGCTGCTCC GGGAGTGGC CGTCAAGGTG AAGGAGGCT ACCGGTTCCG 480

CGTGGCACTG CCTGCTACC CAGCGTGGCT CACGACGTC TCCCTGGCG CTGAGCGAGC 540

50

TGCGCCCCAA GCACTCAGGT ATCTATGCT GTGAGTCCA GCACGCGATC GATGACAGCA 600

GCGACGCTGT GAGGTCAAG GTCAAGGTA TCCATCCAG ACGCCAGAG AGGCTGTTA 660

CGGAGACATG GATGGCTTCC CCGGGGTCCG GAACTATGCT GTGCTGGACC CGGATGACCT 720

55

CTATGATGTG TACTGTTATG CTGAAGACCT AATGAGAGAA CTGTTCTCTG GTGACCTCC 780

AGAGAAAGTG ACATGAGG AAGCACGCGC GTACTGCCAG GAGCGGGGTG CAGAGATTG 840

60

CACCACGGGC CAAGTATG CAGCCTGGA TGGTGGCTG GACCACTGCA GCCCAGGGTG 900

200

GCTAGCTGAT GGCATCTGTG GTACCCCAT CCGACACCT AGCCATGGCT GTGCTTBAAG 960
 CTGCTCTGCT GCGAGAGCTG TTTTCTCTTT CCGCAACCA AGTGGCTTCC CCATAAGCA 1020
 5 CAGGCTCTTC AATCTCTACT GTTTCGAGA CTGGCCCAAG CTCTTCCCAT CCGTGAGGG 1080
 TCGAATCGAG CCGCAACCT AGCTTGTATG GATTAGAGGG TATCTTCACA GTGACAGAGA 1140
 CCGTGAGAGA AGTGCAGCTG CCGCAGGAAG CCAGAGAGAG TGAATCCCGT GGGGCCATCT 1200
 10 ACTCCATCCC CATCATGGAG GACGGAGGAG GTGGAAGCTC CACTCCAGAA GACCCAGCA 1260
 AGGCCCTTAG GAGCTCTTA GAATTGAAA CACAATCCAT GGTACCGCCC ACGGGTTC 1320
 15 CAGAGAGGA AGTAAAGCA TTGGAGGAAG AAGAGAAATA TGAAGATGAA GAAGAGAAA 1380
 AGGAGAGAGA AGAGAGAGAG GAGGTGAGG ATGATGCTCT GTGGCATGG CCCAGGAG 1440
 TCAGCAGCTT GACCCCTGAG GCTCTCTCT CCACTGAGCC AGCAGCCCAG GAGGACTCA 1500
 20 TCTCCAGGT GGTAGCAAGG GCACTCTCT AGGTGCTTC ATCACCCTT CCGATGGAG 1560
 AGTCAGAAC TTTAGGCTT CCAAGGCTC ATGACCACC TACTGAGACT CTGCCACT 1620
 25 CCAGGAGAG GAACCTAGCA TCCCCATCA CTCTACTCT GGTGAGGCA AGAGAGGTG 1680
 GGGAGGCAAC TGTGGTCTT GAGCTATCTG GGTCCCTGA 1720

30

(2) INFORMATION FOR SEQ ID NO: 54:

35 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 1117 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40 (ii) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

GGACAGAGCT CAACCTTCGG GCGGTGAGG CCGCGGCCGA GAGCGGCGG ACTCCCGGG 60
 CCGGAGCTCG AGGCATTGCG GCGTGGCTT CCGAGCTAC CGAGGGCCTG AGCCTTGA 120
 45 GCAGNAGGAG GCGAGGAGAG AGTGCGGCTC CTCTATCGGG ACCCCCTCCC CATGTGGAT 180
 TCTCCAGCTG GGTGGGCTG AGTGGGCTC CCGAGGATG ATGAGGCTTC GCGGCTCT 240
 50 GGTGGGCTC CTCTGGGCTC TCTGGCTCT CTCTGGGACC CCGTGGATG CATTGAGT 300
 TCGAGATGCT TATGAACCTT GTGTAATGA AATATGTGT GTTACCTACC ACAATAGCA 360
 AGGATACTTC AAAGTCCAG AAGGCTTCTT GAGGAAATAT TCTCAACATC GAGATCCT 420
 55 TGAGAGAGAC CGCTGCCAGA ATGGTGGGAC TGTGTGGGC CAGGTCATGC TGGGAAAG 480
 CACGTGCGCA TGTGGCTCAG GGTTCACAG AGAGGACTTC CAGTACTCGA CATCTCAT 540
 60 ATGCTTTGTG TCTCGACCTT GCGTGAATGG CCGACATGC CATATGCTCA GCGGGGAT 600

CTATGAGTGC ACCTGTCAAG TGGGTTTAC AGSTAAGGAG TGCCAATGSA CCGATGCCTG 660
 CCTGTCTCAT CCTGTGCAA ATGGAAGTAC CTTTACCACT GTGGCCAACC ACTTCTTGCA 720
 5 AATGCTTCAC AGGCTTCACA GGGCAGAAAT GTGAGACTGA TGTCAATGAG TGTGACATTC 780
 CAGGACACTG CCAGCATGGT GGCACCTGDC TCAACCTGCC TGGTTCTTAC CAGTGTACAG 840
 10 GCCTTCAGGG CTTCACAGGC CACTACTGTG ACAGCCTGTA TGTGCCCTGT GCACCTCTGT 900
 CTTGTCTCAA TGGAGGCACC TGTGGGCAGA CTGGTGACTT CACTTTTGAG TGCAACTGCC 960
 TTCCAGAAAC AGTGAGAAGA GGAACAGAGT TGTGGGAAAG AGACAGGGAA GTCTGGAATG 1020
 15 GAAAGAACA CGATGAGAAT TAGACACTG AATATATGTA TGTGTGCTTA ATAAAGTGCT 1080
 TTAAACTGAA AAAAAAAAAA AAAAAAAAAA AAAAAA 1117

20

(2) INFORMATION FOR SEQ ID NO: 55:

25

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1903 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

GGCACGAGCT CCGAGAGGCG GCGCCCTGSA GTAGGCCAGG AGCCTCTCTT GCAACTTCTG 60
 35 CCACCGCGGG CCACCGCGGC CGCTGATCC CGCAGAGGAA GGTGCGGGCC GTGGAGCGAT 120
 GACCGCGGCG GGTCCGGGCG GGGCCCGGEG GTTGCCACAG CCGCCGCCGC TTCTGCTGCT 180
 40 GCTGCTGCTG CCGCTGTGTG TAGTCACGCG GGAGCCGCCG AAACCTGCAG GAGTCTACTA 240
 TGCAACTGCA TACTGGATGC CTGCTGAAAA GACAGTACAA GTCAAAAATG TAATGGACAA 300
 GAATGGGGAC GCCTATGGCT TTTACAATAA CTCTGTGAAA ACCACAGGCT GGGGCATCCT 360
 45 GGAGATCAGA GCTGGCTATG GCTCTCAAAC CTTGAGCAAT GAGATCATCA TGTGTGTGGC 420
 TGCTCTTTTG GAGGGTTACC TCATTGCCCC ACACATGAAT GACCACTACA CAAACCTCTA 480
 CCCACAGCTG ATCAGGAAAC CTTCATCATG GATAAAGTG CAGGATTTTA TGGAGAAGCA 540
 50 AGATAAGGTG GACCCGGAAA AATATCAAAG AATACAAGAC TGATTATTTT TGGAGACATA 600
 CAGGCTATGT GATGGCACA AATAGTGGGC TGTATGTAGG AGCAAAGAAG AGGCTATAT 660
 55 TAGAAGGGAC AAAGCCAATG ACCCTGTTCG AGATTCAGTT CTTGAATAGT GTTGGAGATC 720
 TATTGATCT GATTCCTTCA CTCTCTTCCA CAAAAACGG CAGCCTAAG GTTTTAAAGA 780
 GATGGGACAT GGGACATTGC TCCGTCTTTA TCAAGGTTCT TCCTGGATTG GAGAACATCC 840
 60

	TTTTGGCTCA CTGAAGCTGG TACAGGTATG CAGTCATGCT CAGGATATAT AAAGACTGGG	960
	ACTTCAACAT CATAGATAAA GATACCAACA CTAGTGGGCT GTTTTTCAGC ACTTACCCAG	965
5	GGTTTTTNSA GTCTCTGGAT GATTNTIACA TTCTTAGCAG TGSATTGATA TTGCTSCAGA	1020
	CCACAAACAG TGTCTTAAAT AAAACCCGCG TAAAGCAGGT AATACCCGAG ACTTCTCTGT	1080
10	CGTGGCAAG AGTCCCTGTC GTCATATGA TGGCAGATAG TGGCAAGAGG TGGGCAGACA	1140
	TTTTTDCAAA ATACAACCTC GGCACCTATA ACAATCAATA CATGCTTCTG GAATCGAAGA	1200
	AAGTAAAGCT GAACACAGT CTGACAAAAG GCACTCTGTA CATTCTGGAG CAAATTCCTA	1260
15	CATATGTAGA ATATCTGSA CAACTGATG TTCTAGGSA AGGATATGCG CCGTCTTACA	1320
	ATCTTCTCTT CCATGAAAAA AATACAACT GAGTGGCTA TCCACTCTTA GTTCAGAAGC	1380
20	TGGGCTTNSA CTACTCTTAT GATTTAGCTC CACGAGCCAA AATTTTCGCG CGTCACCAAG	1440
	GSAAATGAC TGATACGGCA TCTATGAAAT ATATCATGCG ATACACAAT TATAAGAAGG	1500
	ATCCTTACAG TAGAGGTGAC CCGTCTAATA CCACTCTGCTG CCGTCAAGGAC CCGCAACTCA	1560
25	CGTAACCCAA GTCTCTGAG GTCTCTTATGA CACAAAAGGT GCTAGATATY TACCTAGCAT	1620
	CTCAGTACAC ATCTCTATGCG ATAAGTGTG CCACAGTACA AGTGGGCTC CCGTCTTTTC	1680
30	GCTGGGACCG TTTCAACAAA ACTCTACATC AGGGCATGCG AGAGGTCTAC AATTTTGATT	1740
	TTATTACAT GAAACCAATT TCGAACTCG ATATAAAATG AAGGAGGGAG ATGACGGACT	1800
	AGAAGACTGT AATAAGATA CCAAAAGGAC TATTTTAGCT ATTTTTTTTC CATCAGAATT	1860
35	ATGCAATAAA ATATATTAAAT TTCTCAAAAA AAAAAAAAAA AAA	1900

40 (2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1869 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

50	ACAGCTTTTC GGGGCGGAG CCGCAGCAG CCAAGAGAGG GGGGCGGSA CAGCTCGAA	60
	CTGCGGCGCG CTGCGGCTTC CCGGCTTCG CTGCTCTGCG CCGCTGGGG TGGGCGGCGC	120
55	AGGATGCTGC AGGGGCTTGG CTGCTCTGCTG CTGCTCTTCC TGGCTTGGCA CTGCTGCTG	180
	GCTCTGGGCG GGGGCTCTT CCGCTTGGC CAGCCCGACT TCTCTTACAA GCGCANCAAT	240
	TGCAAGGCCA TCCCGGTCAA CTTGCACTG TGGCAGGCA TCGAATACCA GAACATGCGG	300
60	CTGCCCAACC TGCTGGGCA CAGAGCATG AAGGAGGTGC TGGAGCAGGC CGGCGCTTGC	360

ATCCCGCTGG TCATGAAGCA GTGCCACCGG GACACCAAGA AGTTCCCTGTG CTCGCTCTTN 420
 5 GCCCCGGTCT GCTCGATGA CCTAGACGAG ACCATCCAGC CATGCCACTC GCTCTGGCTA 480
 CAGGTAAAG ACCGCTGGC CCCGTCATG TCCGCTTCG GYTTCCCTG GCCCGACATG 540
 CTTGATGTG AGCTTTTCC CAGGACAAC GACTTTGCA TCCCCCTCG TAGCAGCGA 600
 10 CACCTCTTC CAGCCACCGA GGAAGCTCCA AAGTATGTG AAGCCTGCA AAATAAAAA 660
 GATGATGACA ACACATAAT GAAAACGCTT TGTAAAAATG ATTTTGCCT GAAATAAAAA 720
 GTGAAGGAGA TAACCTACAT CAACCGAGAT ACCAAAATCA TCCTGGAGAC CAAGAGCAAG 780
 15 ACCATTTACA AGTTGAACGG TGTGTCCGAA AGGGACCTGA AGAAATCGGT GCTGTGGCT 840
 AAAGACAGT TGTAGTGCAC CTGTGAGGAG ATGAACGACA TCAACGCGCC CTATCTGGT 900
 20 ATGGGACAGA AACAGGGTGG GGAGCTGGTG ATCACCTCGG TGAAGCGGTG GCAGAAGGGG 960
 CAGAGAGAGT TCAAGCGCAT CTCGCGCAGC ATCCGCAAGC TGCAGTGCTA GTCCCGGCAT 1020
 CCTGATGGCT CCGACAGGCC TGCTCCAGAG CACGCTGAC CATTTCTGCT CCGGGATCT 1080
 25 AGCTCCCGTT CCCCAAGTAC ACTCCTAGCT GCTCCAGTCT CAGCCTGGGC AGCTTCCCC 1140
 TGCCCTTTTC AGCTTTGCAT CCCCAGCATT TCCGAGTTA TAAGGCCACA GGAGTGGATA 1200
 30 GCTGTTTCA CCTAAGGAA AAGCCACCC GAATCTTGTG GAAATATTCA AACTAATAAA 1260
 ATCATGAATA TTTTATGAA GTTAAAAAT AGCTCACTTT AAAGCTAGTT TTGAATAGG 1320
 35 GCAACTGTGA CTGGGTCTG GTTGGTTGTT GTTGTGTGTT TTGAGTCAGC TGATTTTCA 1380
 TTCCCACTGA GGTGTGATA ACATGCAAT TGCTTCAATT TTCTCTGTGG CCCAAACTTG 1440
 TGGTCCACAA ACCCTGTTGA GATAAAGCTG GCTGTTATCT CAACATCTTC ATCAGCTCCA 1500
 40 GACTGAGACT CAGTGTCTAA GTCTTACAAC AATTCATCAT TTTATACCTT CAATGGGAAC 1560
 TTAAACTGTT ACATGTATCA CATCCAGCT ACAATACTTC CATTTATTAG AAGCACATTA 1620
 ACCATTTCTA TAGCATGATT TCTTCAAGTA AAAGGCAAAA GATATAAATT TTATAATTGA 1680
 45 CTTGAGTACT TTAAGCCTG TTTAAACAT TTCTTACTTA ACTTTTGCAA ATTAAACCCA 1740
 TTGTAGCTTA CCTGTAATAT ACATAGTAGT TTACCTTTAA AAGTTGTAAA AATATTGCT 1800
 50 TAACCAACAC TGTAAATATT TCAGATAAAC ATTATATTCT TGTATATAAA CTTTACATCC 1860
 TGTTTTAC 1869

55

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

60

(A) LENGTH: 1259 base pair:

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ACCGTGTGTC TGAGGAGAG GGTGTGTAG GGTGAGGAG CTGGGCTGCG TGTGGGTG 60
GAATGAGAGC GAGGAGGTC GGTGTGTAG GGTGAGGAG CTGGGCTGCG TGTGGGTG 120
10 CGGTGTGTCG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 180
GTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG 240
15 CGGTGTGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG 300
GAATGAGAGC GAGGAGGTC GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 360
GCGTGTGTCG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 420
20 GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 480
CGGTGTGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG TTTTATGTCG 540
25 ATATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC 600
GAATGAGAGC GAGGAGGTC GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 660
TTTGTGTAGC GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 720
30 GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 780
GAGTGTGTCG TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC 840
35 CAATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC 900
TGTGTGTAGC GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 960
ATATGAGAGC ACCAGGAGC GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 1020
40 GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 1080
GGTGTGTAG TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC 1140
45 ATTGTGTAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC TATGAGAGC 1200
AAAAATGTCG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG GGTGTGTAG 1259

50

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 1186 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

CGGCATGAG AATGCTCTCG CTCTCTCTTC AGCTGCGGT GTTGGGCGG GCGCTGGCG 60
 CCGCAGCCCT CCGACTGATT TCCATGCTG CATTACAAAC TCCACAAA ATGCCAGCA 120
 5 TCCATGACA TGAAGAAGAG AAATTCCTCT TAAATGCCAA AGGDCAGAAA GAAACTTTAT 180
 CCAGCATATG GGAATCACTT ACCAAACAAC TTCTCTCTCT TCTGCTCTCA TACAATGAAG 240
 10 AAAAACTGTT GCTCTGATG ATGATGAAG CTCTGAGTA TCTAGAGAAG AGACAGAAAC 300
 GAGATCTGTC GTTCACTTAT GAAGTGATAG TAGTTGATGA TGGCAGTAA GATCAGACCT 360
 CAAAGGTAGC TTCTAAATAT TGCCAGAAAT ATGGAAGTGA CAAAGTACGT GTGATAACCC 420
 15 TGGTGAAGAA TCGTGAAGAA GGTGAGGCGA TTAGAATGGG TATATTCAGT TCTCGAGGAG 480
 AAAAGATCCT TATGGCAGAT GCTGATGGAG CCACAAAGTT TCCAGATGTT GAGAAATTAG 540
 20 AAAAGGGGCT AAATGATCTA CAGCTTGGC CTAATCAAAT GCTATAGCA TGTGGATCTT 600
 GAGCTCATTT AGAAAAAGAA TCAATGCTC AGCGTCTTA CTCTGCTACT CTCTCATGT 660
 ATGGGTCCA CTCTCTGCTG TGCTCTCTTT GTCTCAAAGG AATCAGGAC ACACAGTGTG 720
 25 GGTTCAAATT ATTTACTCGA GAAGCAGCTT CAGGACGTT TCTATCTCTA CACGTTGAAC 780
 GATGGGCATT TGATGTAGAA CTACTGTACA TAGCACAGTT CTTTAAAATT CCAATAGCAG 840
 30 AAATGCTGT CAACTGACA GAAATGAAG GTCTAAATT AGTTCCATTC TGGAGCTGGT 900
 TACAAATGGG TAAAGACCTA CTTTTTATAC GACTTCGATA TTGACTGGT GCCTGGAGGC 960
 TTGAGCAAAC TCGGAAAATG AATTAGGTTG TTGCTAGTCT TCAAGTGTGT TCTTATGCTT 1020
 35 CAGTGTACA TTTCATTCTA TTGAAACTA AAATTTTAAG TAAAGCTGAA ATAAACTTCT 1080
 TGTCAATGTC TGCTTTTGA TAATTTTAAA GAAATAACTT TCCATAAGTA AAAAATTATA 1140
 40 TATCTCTTTG GATATAAATG ATTTTAAAA GATGTTTATT TAAAAA 1180

45 (2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 428 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59.

55 GATCCCCCGG CTGCAGGATT CGGCACGAGT ACTGATTCTT CACTGAGCTT KGTAGTATA 60
 AGCAGAGTTC CAAGTCTCCC CTAGGCTTGT CTCTACATTT CTTTATCATT CCAGTGGGTA 120
 RGGTTTAGCT GGGGGAAGGA CATTTCATAA GGGTTAGTTG GACTGAGCAG TATGGACATT 180
 60

TGGTCTTTTC AGGAGTACT GTGTTTTTC GTGCTAAGT TGGCTTGGT GGTCTTAAT 24
 TTAATGTGTC AGGATGGGG AAAGGGGGG GGTGTGTGG GAGACTACT TATTATTCT 30
 5 TTTTCTGAC TTAATGTT GTGGTAAT GATAGCTTC TGTTCATCT NTGCACTTC 36
 TTAATAATAA AATTTTTTGA AATTGAAAA AAAAAAAAAA NAAAAACTC GGGGGGGGG 42
 CGGTAC 48
 10

(2) INFORMATION FOR SEQ ID NO: 60:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 501 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

GGTACGAGCT TTGAGCAGGG GACAGCCCGA TTGGGGACAA TGGGTCTCTT TGGGCACAT 60
 25 TTGGTTTTCT GTGGGTCTT CTTACCATG GGTAAAGGAG AAGTCCAAA GGAACAGCA 120
 CGGTCACTT AAGACTACCA GTGCTGAG ATGGAGGCG TGTGATGCG CGGATCTCT 180
 30 TTCATCTG GATCTCTCAT CGTGTGAGC AGAAGATGCG GTTGAAGTT CAACAGCAG 240
 CAGAGGATG GGTAAACCGA TGAAGAGGAG GGAATTCTC GAGCTCAT CCGCTCTCT 300
 TCAACCCGA GGTGTAGAA ACAGCTGAG GATGGAATC GGGCAGGAC TCCCTGGA 360
 35 CCTGACATCT CCAAGCTTC AACTGGGCG GACCGGCGC CTGCGCGGC CTTTCCCA 420
 CCTGCGGC GAGACTGCG CCTGCGGCA AGACTTCCA TAAAACTG GTTCTCTCT 480
 40 AAAAAAAAAA AATAAAAAA A 501

45 (2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1197 base pairs
 (B) TYPE: nucleic acid
 50 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

ACATGATGCT TACCAAGAA TTGTCANAG GCGCGCAGT GAGCAGGTG CTCAATATG 60
 AGTGTCTGAG GAGCTTCTG AGGCGCGCG TGTGTGCTG GGTCTTCCG TACGTGGGG 120
 CCCCCAGGC CCTACCGTG AAGCTGCGAG TGACCAKCAA CAATTTCTT CAGCCACCG 180
 60

	AGAGTGGGGG CCAGGATTTC TTCCAGCGCT GGAAGCAGCT GAGCCTCCCT CAACAGGAGG	240
	CCGAGAAAAT CTTCAAAGCC AATCAACCCA TGGATCCAGA AGTTACTAAG GCGAAGCTTC	300
5	TGGGTTTGG CTCTGCTCTC CTGACAAATG TGGACCCCAA CCGTSAGAAC TTCTGAGGGG	360
	CGGGATCAT CCAGACTAAA GCGCTGCAGG TGGCTGTCTT GTTCGGCTG GAGCCCAATG	420
10	CCGAGGTTCA GATGTACCGG CTGACCTTGG GACCCAGCAA GAGCCCGTTC TCCGTCACG	480
	TCTGTAGAGT GGTGACACAG CATTCTGAG CCGTGGACTC TGCCCCGGGG GATGTGCGG	540
	GCACTGGGCA GCGCCTTGGA CTGAGGCAAT TTCTGTGAT GGGGACCTC CACTGTGAG	600
15	AGAGAAAGACA CCAGGCTTTC GGGATGCTT GGGACTTTCC TCCGGCCTTT TGTATTTTAA	660
	TTTTTGTCA TCTGCTGCTG TTTACATTTT GGGGGTTAG GGGAGTCCC CTTCCCTCCC	720
20	TTCCCCCCCC AAGCACAGAG GGGAGAGGGG CCAGGGAAGT GATGTCTCC TCCCCCCCCA	780
	CCCCAGCTG TTGTAGCCCC TCTACCCCC TCCCCATCCA GGGCTGTGT ATTATGTGA	840
	GCGAATAAAC AGAGAGAGCG TAACAGCCCC ATCTCTGTGT CCATCACCCA CTGTTAGGTA	900
25	GTCAAAGAA GGGGTGAGG GCATGCAGAG TGTGGGTGGC CAGNTTCGCA GCCCATGGGT	960
	GGGACTCTGG GGAGACAGCA GCAGCAGTAG CCGTGAAGC CCGAGCTGCA AGGTACCCAG	1020
30	AGGACTCTCT GTGCTGCTT CCTYAGTCCC CAACACCAGG TAGCAAGCTY TGGGAGCTC	1080
	GGCTGTAGT ACCTCATCTT CTGTCTCTY TGTGTGCCCC GGTCTGTGT GGAAGTGGT	1140
	GGAGGTGACC AGGGIATAGA AGTTTCGGAG CTGATTGAAA GAGGATTAAC TTCCCCG	1197
35		

(2) INFORMATION FOR SEQ ID NO: 62:

- 40 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 595 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

	ATTNANBACK TKYAGCCCTTT WATACMATCA TTATAGGSAAR AAGCTGGTAC GCTGSMARGT	60
50	ACCGGTCTYGS AATTTCGGG TCGACCCACG CTTCCGGTAC AGCGGAGTGT GGTCTGACA	120
	CCAGATCTTC TCTGTCTCTG GTTAATGTCA GTGAGGCTG GAAGTTGAAT AAATGAGAAC	180
55	AGGAGTGGTC TGGGCCCATG TAAATGATCC TCCCTGAAA GAGGAACAG CTTTCATCAT	240
	TTGTTCCAGC TAAGCTTGC ATGCATTATA GATCTGTGTC TAAGCAGTGG GAAAGATCTC	300
	ATAAGTAATG TTTTATGTTT TTTCTGCTC TCTCTCTG TWGTTCTTGG CTTGTGAGT	360
60	GTGTTTGTGT GTTAACTGGA AAATTCCTAT AAGCCAGTTC TCTCTAAGTT TAAAAACGA	420

ATGAGAAAA CCATAAAAATC TCTGGCTAT GCACATGTC CCTGTATTST GAAACATTA 480
 AATGCTAAT AAAAAGGAAG GAGATAGTC ATAAATGTGC ATCAATATA TTCTGAGTK 540
 5 TATAGAAATC ATGACCAAG CATTASACT AGAAGCAAAA AAAAAAAAA AAAAA 595

10
 (2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1478 base pairs
 15 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

20 CGGCGCTGAG GACGCACGGA TGCTTCCTG CCTTCCATC AAGATCTCAA TTTTGTGGC 60
 AASTTCCCTAC AGCCCTCTGT GATGAGAGAG CTGCTCCGG AAGAACCCAG CCAGGATGGA 120
 25 CCCCCTGAATC CGCATGGTGC AAGATTTCC AGCCCTGCAC CAGGCAGCCG AGGACATGAA 180
 GTTCTTTGAT GCCAGTCCCA CTTCTTTGC TTCTCTACTG GGTCAATCC TGGCCATGGA 240
 GGTACTGGCC TGGCTCTTA TTTACCTCT GGTCTCTGG TGGTTCCTCA GTGCCCTGGN 300
 30 CCGTCTTCAT CCTGGCATC TTTAGCTTC AGTCTGGTG TCTGCAGCAT GACCTGGGCC 360
 ATGCTCCATC TTCAAGAAGW CTTCTCTGAA CCACCTGGCC CAGAACTTC TGATGGGGCA 420
 35 GCTAAAGGGC TTCTCCGCC ATTTCTGSA CCTCCGCCAC TTTCAGTACC ACGCCAAGCC 480
 CAACATCTTC CACAAAGACC CAGATCTGAC GGTAGCGCCC GTTTCTCTCC TGGGAGAGTC 540
 ATCTGTCGAG TATGGCAAGA AAAAAAGGAG ATACTACCC TACAACACG ATCACTGTGA 600
 40 CTTCTTCTTG ATGGCCCCG CCGTCTTAC CCTTGTGAAC TTGGAAGTGG AAAATCTGGC 660
 GTACATGCTG GTTTCATGC ATTTCTGGA TTTCTCTTG GGTGCTAGCT TGTATGCCCG 720
 45 CTTCTTCTTA TCTACCTCC CTTCTTACG GTTCTCTGG GTCTCTCTCT TTTTGTGTG 780
 TGTACAGGTC CTGGAAAGCC ACTGCTCTGT GTGATCACA CAGATGACCC ACATCCCCAA 840
 50 GAGATCTGCG CAGAGAAAC ATCGAGATG GGTAGCTCT CAGCTGAGAG CCACTTGCAA 900
 CTTGGAGGCC TCACTTTTCA CCACTGCTT CAGCGGGCAC CCAACTTCT AGATGAGCA 960
 CCACTCTTC CCCAGSATC CAGACACAA CTACAGCCG GTTCTCTCG TGTCTAAGTC 1020
 55 GCTGTGTGCC AAGCACGGCC TCAGCTACGA ATTAAGCCCT TCTCACCAG GCTGTGTGGC 1080
 ATCTCAGGT CCTTAAGAA GTCTCTGAC ATTTGGCTG AGGCTACTT CATTAGTGA 1140
 AGGCAACACC CAGGCGGGCA GAGAAGGGCT CAGGTCACCA GCAACCAAGC CAGCCCCCGG 1200
 60

CGGGATCGAT ACCCTCACCC CTCACCTAGG CAGCCTGGGG GTGCCCCGCC TGCCCCCTCTG 1260
 GTACTGTPTGT CTTCCTCTCG GCCCCCTCAC ATGTGTATTT AGCAGCCCTA TGGCCTTGGC 1320
 5 TCTGGGCTTG ATGGGACAGG GGTAGAGGGA AGGTGAGCAT AGCAGATTTT CCTAGAGCGA 1380
 GAATGGGGGG AAAGCTGTTA TTTTATATT AAAATACATT CAGATGTAAA AAAAAAAAAA 1440
 AAAAACTGGA GGGGGGKCC CGNAACCAA TTGCCCC 1478
 10

(2) INFORMATION FOR SEQ ID NO: 64:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1033 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

20

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

GGCACGAGGA AGAACGAAA GGTGAGAACA TGGACGTAA TATCGCCCCA CTCGCGCCT 60
 25 GGGACGATTT CTTCCTGGGT TCGATCGCT TTGCCCCGCC GGAATTCAGG GACATTTCCA 120
 AATGGAACAA CCGCGTACTG AGCAACCTGC TCTATTACCA GACCAACTAC CTGGTGGTGG 180
 30 CTGCCATGAT GATTTGATT GTGGGTTTC TTAGTCCCTT CAACATGATC CTGGGAGGAA 240
 TCGTGTGTGT GCTGTGTTC AAGGGTTTG TGTGGGCAGC CCACAATAAA GAGTCTCTTC 300
 GTCGGATGAA GAAGCGCTAC CCAAGACGT TCGTTATGGT GGTGATGTTG GCGAGCTATT 360
 35 TCCTTATCTC CATGTTTGGA GAGTCTATG TCTTTGTGTT TGGCATTAAT TTCTCTTTCG 420
 TGTGTATGTT TATCCATGCA TCGTTGAGAC TTCGGAACCT CAAGAACAAA CTGGAGAATA 480
 40 AAATGGAAGG AATAGGTTTG AAGAGGACAC CGATGGGCAT TGTCTGATG GCCCTAGAAC 540
 AGCAGGAAGA AGGCATCAAC AACTCACTG ACTATATCAG CAAAGTGAAG GAATAAACAT 600
 AACTTACCTG AGCTAGGGTT GAGCAGAGAA TTGAGTTGCA GTTTGCCCTT GTCCAGACCT 660
 45 ATGTTCTGCT TGCCTTTTG AAACAAGAGG TGCACGTACC ACCCAATTAT CTATGTCAGC 720
 ATGCATGTAT AGGCCGAAT ATTATCAGCT CTGATGTTTC AGAGAGAAGA CCTCAGAAAC 780
 50 CGAAAGAAAA CCACCACCT CATTATGTT CTGAATTTT AGGTGTGTTT ATGAAATCTA 840
 ATGGGAAATG GATCAGACGA TTTCTTTAAG GGAATTAAAA AAAATAAAG AATTACGGCT 900
 TTTACAGCAA CAATACGATT ATCTTATAGG AAAAAAAAAAT CATTGTAAAG TATCAAGACA 960
 55 ATACGAGTAA ATGAAAAGGC TGTAAAGTA GATGACATCA TGTGTTAGCC GTTCTCTAAT 1020
 CCCCAGAAAT TGTAAATGTT GAGATATAAA TTAGTTTTTA TATTTCTCTT AAAATCAAA 1080
 60 GATGATCTCT ATCACTTTC CACCTTTTG ATGTGACGTG GAAACTGTTT AAGCCAGTTC 1140

TTCTACTTC CTCTAGAAA ATAAAGATAG CTTCTTAGGA TACCTTCTTA CATTCTTTTA 1200
 AATTTTTSAA ATCTAGTAA TGCTCTTCA CAGCAAGTA TCTCTTGCAA ACTTAATGTC 1200
 5 AATTTCTTAA AGATGCTTAA ACTTATGTAA CTTCTATTAI TCTGACGGA CTATTAIAAA 1300
 TACAAACAGA CAAATATTA AAAAARCTT GATTCTATT TACCTTACAC ATCTCTTCTT 1300
 10 GTTACAGTGA AAAAATGCT CCAAGAAAAT GTTCTCATI TTTCTATCTT TTCTTTTCTA 1400
 ACTGAACTT TTAGAAGAA GAAATGAAT GTGCACTTTA TCAATCTCTT AGGGGCACAA 1500
 GGAGACAAI AATAGCTGAT CTTTGAAT TTAAGAAAG TCTTATATG ACCAAGCAAA 1500
 15 AAGCTTTTAA AATCTTAAI GAAATGGA TGCAGCTACT GCACTTATA AAAAATTTTA 1600
 GATAGCAAT CTTAAAGTA TATGCTTTA TACCTAGATA TTAGAATTAT GATAGCATGA 1600
 20 GTTCTATAT TCTATCTT TCTCTCTT TCTCTCTT TTAGAATAG CTAAATAAAA 1700
 ATCTTTTCTC TCTAATTA ATGATCTCT AGCTGAAGTA GAAACATTTA GCTTCTCTA 1800
 GCATTAAAT GTGAAGACA CTGAGCTCT ACTTACTGAA GAACTCTCT CTATGCTCTA 1800
 25 GAATAAGAAG CAATGATCT CTCTCTCTG TTTTCTCTC ATTTTAACT CTGAGCCAAC 1900
 CTACAGCAT GATCTTTAG ACTATATAT CAGCATGCT TCTAGACAT GGTCTAGAA 1900
 30 CTGTACCTT ACCACATTA GAAGAATAAA ATGATTTAA GTTAAAAAA AAA 2000

35 (2) INFORMATION FOR SEQ ID NO: 65:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 440 base pairs
 - (E) TYPE: nucleic acid
 - 40 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 65

45 ATGTTTCTTA CTAGAACTT GTCTCAAC TATATAGCT TAATTTCTT GGTTCATCT 60
 GTGCTCTAG TATCTGGA GGCTCTATG GAGATAGCA GAGAAATAT TTTTCTCT 120
 AATTAATGG TCAATCAT TCTCTCTT TGCTCTAT TATCTCTCT CTATTTGCA 180
 50 CCTCTCTCT TCTAGCTAG TTAATATAA TGTCTTTT TCTCTCTCT TCTCTAGTA 240
 ATGCTTCTG TAAAGAAAT CTCTCTCTT AATTCAGTA TATCTCTATA TCTCTCTT 300
 55 GGCTCTCCA TTTTCACT AGGATCTCT AAAAGATGT TCTTACAGGA TATCTAGAA 360
 AATCCAATG GATCTCTCT CAGCTCTCT CCAGCTCTG AGACAGAGG AGACTCTAT 420
 TCAAAAAAA TTAAAAAAA 440
 60

(2) INFORMATION FOR SEQ ID NO: 66:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3301 base pairs:

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 66:

GGTCATAAGG GGAGGCTTGN NETGTGTCCC TCCAGTTTGT GCAGAGGGGA TTAGAAGTAA 60
 15 GTAGGTTAGA GGGGAGGTGG AGGGAGTGTG CTGGGTTGTG AGCTTTTATG ATGCTGAAAG 120
 GATCATGATA TGCTAAGGAC AGGATAGTGT TGGTTGTAC ACACAGGTGT AGGCAATCCT 180
 20 GGTGGTTAGT ATGIAAAAGT GAATGTCTG ACTCCCTTAG AGGTACCTG NCAGAGTGCC 240
 CTTGGARGGA CTAGTGCTGG AGAAATTAAT AGSAGAGGGG ACGGGCATCC ATTAACCTT 300
 TCTTGCTTGC AGCCTGTAGG GTCCAGCGTC AAAGCGAATC ATGGGGTCCA GGGCTGAGCT 360
 25 GTGCACTCTC TTAGGCGGAT TCTCCTTCTT CTTGCTACTG ATACCAGGCG AGGGGGCCAA 420
 GGGTGGATCC CTCAGAGAGA GTCAGGGAGT CTGCTCCAAG CAGACACTGG TGGTCCCGCT 480
 30 CCACTACAAC GAGTCTTACA GCCAACCAGT GTACAAGCCC TACCTGACCT TGTGCGCTG 540
 GAGCGCATCT GCAGCACTTA CAGGACCATG TACCGCCTTA TGTGGCGGGA GGTGAGGCG 600
 GAGGTTACGC AGACCCATGC AGTGTGCTCC CAGGGCTGGA AGAAGCGGCA CCCGGGGGCG 660
 35 CTCACCTGTG AAGCCATCTG CGCCAAGCCT TGCTGTAACG GAGGCGTCTG CTTAGGCCT 720
 GACCACTGCG ATGTGCGCCC CGGCTGGGGA GGGAAGCACT GTCATGTGGA CGTGGATGAA 780
 40 TGTAGGACCA GATCACCTT CTGCTCGCAC CATGTGTTTA ATACGGCARG CAGCTTCAM 840
 TCGGGCTGCC CCAATGACCTA GTGCTAGCGG TGACCGGGCG CACCTGCATG GAGGGGTCC 900
 CAGAGCCCCC AACCACTGCC AGCATACTCA GCGTGCCCTT TCGGGARGCG GAAAAAGATG 960
 45 ACGCGCTCTG AAGCAGGAGA TTCACGAGCT GCGAGGCCCT TGAAGCGGCT GGAGCAGTGA 1020
 NCCGGTCAGC TGGGCCCTGG NTCAGAGGT GTTGCCCGTG CCGCCTGAAG WGTGTCAGCT 1080
 50 AGAACAGSTG GGTGAGCTGT GGGGCCGCGG TGACCGGATC GAATCTCTCA GCGACCAGGT 1140
 GCTGCTGTG GAGGAGAGGC TAGGTGCTG CTCTGTGAG GACAACAGCC TGGGCCTCGG 1200
 CGTCAATCAT CATAGAAG CCTCTACAGC ACGCTGCCC CTAATTTAT ACAGAAACCG 1260
 55 GACCCACTAA TCTCTGGA TTGGCCGACT GTGAGGTGCA GATAAGGCTA TCAGCCACCT 1320
 AAGAGCAATG AACAAAGGAA ACTTCAGAGA GTGAAGAAA GGGGAGGCC TGTGTTCTTC 1380
 60 GCCTSCCCCT GATCTTCTG GCTGGGGGCA GGTGCGCTGG GCAAGAAGTG CTTCTTCAAT 1440

	TCCTTAACAA ATGCAACCAC CAACATCCAG AATCTCTCTCT CTCTTTATTT TTAGTTTMTT	1500
5	TTCTGTATATC CAGATAATTA AAAAAACCA A LAAGLAAA ACTGGGTCCC ACCCTCTCTT	1560
	TTTGGTCCA GCTACCTCC CCAATTCTGT GACAGGTCT GAGTGAAG GAGGAGTCT	1620
	GTAATGCON CAGGAAGAA ATGAAACTG TTTCAGAGAG GGGGAAGCT CAACGAAA	1680
10	AGAAATAAAT TAAAAGCCCT CCTATCCCT CCAGCTAGGG TTGGTTCCTT TCCCCAATC	1740
	CCCAGGCGG CAAAGTGAAT GCAGCAGCTG AATCTGTCTT CTTCCTCTTG TGTCTGCTGA	1800
15	GATGCTCAG CAGGGCTGCA GGGGCTGGG TCTCTCATG TCACTGAAG AACTGTACTA	1860
	TGGGACAGA AAACAGAAA TGTGAGACT GAATCTAT CCAAGAGAT GTATGACCTT	1920
	GGGATCTGG GCAAGGGCAG GATGAGACT TCTGAATTAG AAGGCTCCAG CCCCCACTGA	1980
20	CAGGAGCTA CACTGGGAG GAAGGTAAAG CTCTGAGGA AAGCTCCAT GATGAGCTT	2040
	GGAGTCTTC AGGTATCAGC TTCTAGTCA AGCTGAGAA GTCTCTCTCA CAATCTGAT	2100
25	AGTCCATTGA ATCCATGGAC TTTGAGTGG GGGGATTTG TTCCAAAGAA TGATGAGT	2160
	CACTGCCCCA TCTGGGTAG AGGGGTAGAG AAGACCAT AGAAGAGAC TCACTGGG	2220
	ATGGAATGTT CCCCCTCCTT GTCTAGCTG AGTCACTGA GATGAGGGG AGGCAACTGT	2280
30	CCCACAGACA AATCAGTAG AGGTGGGCT CAAGAGTGA GACTGCAAG AGGCAAGAT	2340
	CCATGATGG GGCAGAGAG GGCAGAGT GGGGCTGAT CCACATTTCA CTTCAGAAGT	2400
35	TGAAGATTC AAAGAGGAGA AATAGTGGG AGAGGGGAGA CAAGGAAGAG GATTTTGGC	2460
	TGCTTCAGG CCACTGGGT GGTAGCTGT GGTGAGGAAG ATGGGACAG ATGGGAGGA	2520
	AGTCAGAGC CAGGGTTCAC CCACGCTCC CAATTTCTT CATATAGTCA CCACCACT	2580
40	GGCATCAAT GAGATTTC CGAAGACAG TGAAGCATG AGTGGCGAC TCTGTCAGC	2640
	AGAGTGGGA CTTCATCTG TGTAGCTCT TGTGAGGA CTGGGGCAG CACCGGCACT	2700
45	TGACATCTC CAGAGTGAA GGAAGCTCT TCTTGCAAT AGAAGCTTT GTAGAGGA	2760
	ATGACTATG GACCAATGG AACCTGCTG TGAATGCA GATGAGAGT GGCAGTTT	2820
	AAAGGAGAA TGAAGCATT GTGTCTCTT GTTCTCTAG GAAACACCA GAATCTCAG	2880
50	TCCTGATGA GAGGCTCT GGTAGCTCA GATATGAC TGAATGCA CTCTCTCT	2940
	TTCTGCTGA TGAAGATGAC TCTGCTAGT CAGGAGGCA GTCGAGAG CAGGCTCAG	3000
55	GTATGCTCT TGGCAATGG GACAGAGCT GATGAGAGT AATCTCATCA TCCAGAGCA	3060
	ATGAGAGAG CTCTGCTGT TGAAGACAA AATATAGGC TCGAGGGAG GGAATGCT	3120
	AGCCCTCTG AATCTCACT GATCTCTGA CATATTTT ATGAGGAG GATATGAGC	3180
60	CAAGATCTT ATCTCTCTGA CTCTGCTCT TGTAGGCA CACACAGGA TGGGAGCA	3240

AGTCAGGAAG AGGATCCAGC CATTCGTAGAA GGCCATCTTG AAGAAGTACT TGSCACTGGG 3300
 5 G 3301

10 (2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1535 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

GGTAAGAGGT CAAGCGAAG GATTTCAAGS AACAGATCAT CCACCATGTG TTCACCATCA 60
 20 TTTCATCAG CTTTCCTGG TTTGCCAATT ACATCCGAGC TGGACTTA ATCATGGCTC 120
 TGATGGACT CTTCGGATTA CCTGCTGAG TCAGCCAAGA TGTTTAACTA CGCGGATGG 180
 25 AAGAACACCT GCAACAACAT CTTCATGCTC TTGGCATTG TTTTATCAT CACCGACTG 240
 GTATCTCTGC CCTCTGGAT CCTGATTTG ACCCTGGTGT ACCCACTGSA GCTCTATCT 300
 GCTTCTTTG GCIATTACTT CTTCAAATTC ATGATGGGAG TTCTACAGCT GTTGTATATC 360
 30 TTCTGCGCCT ACCTCATTTT GGGATGACC CACAAGTTCA TAACTGAAA GCTGCTAGAA 420
 GATGAACGCA GTACCGGAA GAAACAGAGA GTTCAGAGGG GSAGGAGGCT GATGCTGGG 480
 35 GATGAGCAAA GAGCGGCCC CTAGTCAATG GCCACCCCAT CCTCAATAAC AACCATCGTA 540
 AGAATGACTG AACCATTTT CCAGCTGCT CCCAGATTAA TGCATAAAG CAAGGAACTA 600
 CCGCGCTCCC TGGCTATAG GTTCACTTTA AGCTCTGGG AAAAAGGAGA AAGTGAGAGG 660
 40 AGAGTTCTCT GCATCTCCC TCTTGTCTG TCACCCAGTT GCCTTTAAAC CAAATTTCTA 720
 CCAGCTATC CCCAGGTAGG GGAAGTTG TATATTTCTG TTAGAGGGGG AGGTCTGTAT 780
 45 TTCTCTCCT ACCCGCCAAG TCATCTTTT TACTGCTTTT GAGGCTCTCT CTCAGCTCTC 840
 TGTGCTAGG GGTACAATT CACATCTCT ATTCTGAGAA TTTGGGCTCA GTGTTTGCC 900
 50 TTGACTCCC TGACCTCCAG AGCCAGGCT GTGCTTTATT GTCCATCTG TGGGCTCAT 960
 TCTGCCAAAG CTGGACCAAG GCTAAGCTTT CTAAGCTCCC TAACTTGGG CAGAAACCAA 1020
 AGCTGAGCTT TTAAGTTCT CCTCTATGA CACAAATGAA TTGAAGGTA GAGGAGGCT 1080
 55 CAGATAACCC TTACCTTACC TCTGCAAAA AGTGGGGGCT GTACTGAGGA CTGCTGGAT 1140
 GATCTTTCTT AGTGTACTT CTTTCACTG TCCCTGTAGC GACAGCTTA AGATCTGACT 1200
 60 GCTCTCTCT TTCTCTGCGC TCTTCTCTT TCCCTTTCT CTTCAAGTAG GCTAGCTGGT 1260

214

TTGGAGTAGA ATGGGAACIA ATTCTAATT TTATTATIA AACATTGCG GTTTTGSTTT 1320
 TAAAGCCAGA ATTAGGATA GCACTAGCA TTTCAGCAGA GGGACCATTT TAGACCAAAA 1380
 5 TGTACTGTTA ATGGSTTTT TTCTAAATT AAAGATTAA ATAAAAATA TTAATAAAA 1440
 CATGCCATA AGTGTAGAG TATTAGGAT TGAGAGGGG GATCACTAA ATAAACGAAC 1500
 ASAGCTTTC TTATGAAAA AAAAAAAAAA AAAA 1560
 10

15 (2) INFORMATION FOR SEQ II NO: 68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1244 base pairs
- (E) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 68:

25 GGGCACCCAC CAGCGGAGGC GACCTCAGCG CGCACTATG GGCTGGCTAC CAGGACATGC 60
 GGAGACTGGT GCACGACCTC CAGCCCCCG AGGTCTGCAG TCTCTGAAC CCAGCAGCCA 120
 TCTACGCCAA CAACGAGATC AGCTTGCCTG ACCTTGAGCT CTACGGCTTT GACTACGACT 180
 30 AACCCCTGGC CCAGTATGCA GATCACTGC ACCCGAGAT CTTCAGTACC GTCCTGACA 240
 TCTGATCGA GCACTACAAG TACTCAGAAG GATTCGGAA GTATGACTAC AACCCAGCT 300
 TTGTCATCG TGGCTTCAG TATGACATC AGAAGAGCCT TCTGATGAAG ATTGACGCC 360
 35 TCCACTAGCT GCAGTGGAG ACAACCTACA GGGGCTCCA GCTCTGCA GAGGAGGAG 420
 TATGAGCT GTATGGGCT ACACGACA TCCACTATA CCAGATGAGT GCTTCTATG 480
 40 GCAAGGCTCC CTCATTAAAG CATTTCATGG ACATCTTCTC GTACCGAG AGTGGCTCTGC 540
 TCTCTCTAT GGTGACTAC TTTTGGGTC ACAGCTGGA GTTGACCAA GACATCTCT 600
 ACAAGGAGCT GACGAGGCT ATTGAGAGC TGCATGTGAA GGGCTCATG TACAGTGA 660
 45 TCGAGCAGGA CATGAGAAAG TACATCTGA GAGGAGATGA GAGTCTGCT GTCTGAGCC 720
 GCTCTCTGC CCATGGAAG CATTCTGC TATACCAA CATTCTCTC ATTCTGTAG 780
 50 ACAAGGAGAT GGGGACAGT GGTCTCTGC ATTGAGGCA CTTCTGATG TCTCTATGT 840
 CCAGGAGAC AAGCCAGCT TCTCACTGA CCGGCGAAG CTTTACAAA AACTGATGA 900
 GAAGGCTCA CTTCAGTGG AGGATACAC CGCTTGAAG AAGGGAAGA TATATCGCA 960
 55 GCGAAACCTG TTTGACTTCT TACCTTAC GGAATGGCT GGGCTGCG TCTCTACTT 1020
 CCGGACACAC CTCTATAGT ATCTGGGGA TCTCATGCT GCGCAGGCT GCGGACAGG 1080
 60 CGCCATCATC CCGAGCTCG ACGTGAGAT CCGCATCAT AACACGGAGC AGTACATGA 1140

CTCGCTKACG TGGCAGCAGG CGCTCAGGGG GCTKCTKGAG CGCATKACA CCTATCAGGA 1200
 CGCGGAGTTG AGGCAGGTCT TGGTTCCTTG ATGAAAGANC GRT 1244

5

10

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1292 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

20

GGCAGGAGCA GCGACGCGAC TGTGCTGCGG GCCGTCTTCT TCCCCCGAG CTGGGCGTGC 60

GCGGCCGCAA TGAAGTGGGA GTTGCTGCTG TGGCTGCTGG TGTGTGCGC GCTGCTCCTG 120

CTCTTGCTGC AGCTGTGCG CTTCCTGAGG GCTGACGGCG ACCTGACGCT ACTATGGGCC 180

25

GAGTGGCAGG GACGACGCTC AGAATGCGAG CTGACTGATA TGTGCTGCTG GGTGACTGGA 240

GCTTCGAGTG GAATGCTGGA GAGGTGCTT TACCAGTTGT CTAAACTAGG AGTTTCTCTT 300

GTGCTGCTAG CCGAAGAGAT GATGAGCTG GAAAGGGTGA AAAGAAGATG CCTAGAGAAT 360

30

GGCAATTTAA AAGAAAAAGA TATACTTGTT TTGCCCTTG ACCTGACCGA CACTGGTTCC 420

CATGAAGCGG CTACCAAAGC TTTCTCTCAG GAGTTTGCTA GAATCGACAT TGTGGTCAAC 480

35

AATGCTGGAA TGTCCAGCG TTCTCTGTGC ATGGATACCA GTTGGATGT CTACAGAAAG 540

CTAATAGAGC TTAAGTACTT AGGAGGCTG TCCTTGACAA AATGTCTCTT GCTTCACATG 600

ATCGAGAGGA AGCAAGGAAA GATTGTTACT GTGAATAGCA TCTGCGGTAT CATATCTGTA 660

40

CCTCTTTCCA TTGATACTG TGTAGCAAG CATGCTCTCC GGGTTTTTTT TAATGGCCTT 720

CGAACAGAAC TTGTCACATA CCCAGTATA ATAGTTCTA ACATTTGCC AGGACCTGTG 780

45

CAATCAATA TTGTGAGAA TTCCCTAGCT GGAGAAGTCA CAAAGACTAT AGGCAATAAT 840

GGAGACCAAT CCCACAAGAT GACAACCACT CGTTGTGTGC GGTGATGTT AATCAGCATG 900

GCCAATGATT TGAAGAAGT TTGATCTCA GAACAACCTT TCTTCTTAG TAACATATT 960

50

GTGGCAATAC ATGCCAATCT GGGCTGCTG GATAACCAAC AAGATGCGGA AGAAAAGGAT 1020

TGAGAACTTT AAGAGTGTG TGTAGCAGA CTCTCTTAT TTTAAATCT TTAAGACAAA 1080

55

ACATGACTGA AAAGAGCACT TGTACTTTTC AAGCCACTGG AAGGAGAAAT GGAAAACATG 1140

AAAACAGCAA TCTTCTTAG CTCTGATA ATCAAAGACT AATTCTGAT TTTACTTTTT 1200

AATAGATATG ACTTGTCTT CAACATGGAA TGAAATAAAA AATAAATAAT AAAAGATTGC 1260

60

CTGGAATCTT GCGAAAAA AAAAAAAAAA A

1291

5

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 1031 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:

15

CGGCTGTGCG TTTCGAGGAG AACGATATAT TACTCTCTCT GGAATGAGGT TTCTGCAGG 63
 CATTGTATG TAAAGAGAG AGTATCTCTT CAGGTGAGCA GTTTTGAGGA CCGTATGTG 117
 20 GGAATTTTA AGTATATA TTAACATCA TCCACTGGG AGGTGTTATG TTGTATCCC 181
 ACTTTGGGTC GTTCTAGTTT TTCTCTAGC CTAGAGGAG TTCTCTGTC GAGGCTGGT 245
 CTCTTGCTTA CCGCTTCA TGAACAGGG GAGGAATATT TACTTCCCA CCGCTTGTG 309
 25 TTTCTCTTCA CTATACAC TGAATGGAAC TGTCTCTGT ACTCTCTGT CTGCGGATT 363
 ATCTCCGAG ACTTAGCT GTTCTAGTG AGCTGAGAT CTGACAAACA GTCTATGTC 427
 30 ATGCAATGAA GAAATGCT TGCACAGC AGAGGGAACA CTACAGCCC AGGCGCTTT 481
 ATTTTGAGAA AGGTCTCTG GGCCTCTTAC TTCTCTCTCT GTTATAAAG CAGATATGT 545
 GGCATCTTT CCGAGCTTA GAGTGGGCTC CTCTCTTTT GGAATCTTT TCTCTCTCT 609
 35 TGTCTAGTGC TCTCTCTCT CAGGCTTCT GGCACAGG TCTCTGCTGT GTTCTGAGT 663
 GAGTGTGCT GAGGCTTT TTCTCTCTCT TCTCTGAAAG AGAGGCTCT GCGATAGAG 727
 40 ATGAGAAACA ACAGCTCTC CTTCAGACAA TGAGGCTATC TCTCTCTCTG CTGCCATCT 781
 CTCTCTCTAC TGAAGTAC AGCTGTAGG AGCTGAGTCT CAGAGGATT CTGCATTCT 845
 CTACTCTTAG TTCTCTCT CTGATCTCT CCGCTCTCT TCGGCTCTCT TCGCTCTCT 909
 45 GGCATCTTA CCGCTCTCT GGCCTCTTT ACTACAGGC TATCTCTCT GACTGTCTAT 963
 GACTTCTCT CAGAGTCTG GAGCTCTCT CTGAAATAA AGCAAGTAT TTAATAAAA 1027
 50 AAAAAAAAAA A 1031

55

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

60

- (A) LENGTH: 855 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

5 AGCTATTGAC ACTTCTTNGT GBSATCCGAG TGAGGCGACG GGGTAGGGGT TGGCGCTCAG 60
 GCGGCGACCA TGGGCTATCA GGGCCTCACT GGGCTCTCA TBTGATGAG CGTGTCTG 120
 GGGTCTGTGG GTTCTTNGT GCTTGGGTC ATGCTAAGG CTCTAAGCG GGGAGTTAT 180
 10 ATTAACCATGT TBTGACCTG TTAGCTTGC TGCTATCTCI TTTGGCTGAT TGCAATTCTG 240
 GCCCAACTCA AGCTCTCTTT TGSACCGCAA TGGAAAAATG AAACCATCTG GTATCTGAAG 300
 15 TATCATTGSC CTTGAGGAAG AAGACATGCT CTACAGTGCT CACTCTTTGA GGTCAAGAGA 360
 AGAGAATGCC TTGTAGATGC AAAATCACCT CCAAACCAGA CCACTTTTCT TGAAGTGCCT 420
 GTTTTGGCCA TTAGCTGCT TAAAGGTAA CAGCACATTT GAATGCCTTA TTCTACAATG 480
 20 CAGCGTGTIT TCTTTGCT TTTTGCCT TGGTGAATG ACCTGCCTCC ATAACCTGAA 540
 CTGTGCCGAC TGCACAAAAC GATTATGTAC TCTCTGAGA TAGAAGATGC TGTCTTCTG 600
 25 AGAGATACGT TACTCTCTCC TTGGAATCTG TGSATTTGAA GATGGCTCCT GCCTTCTCAC 660
 GTGGGAATCA GTSAGTGT TAGAACTGC TGGAAGACAA ACAAGACTCC AGTGGGGTGA 720
 TCACTAGGAG AGTACCTCA GAGGGAAGAG CCATCTCAAC AGAATCGCAC CAACTATA 780
 30 TTTCAGSATG AATTTCTTCT TTCTGCCATC TTTTGAATA AATATTTTCC TCCTTCTAW 840
 KFAAAAAAAAA ANANI 855

35

(2) INFORMATION FOR SEQ ID NO: 72:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1274 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 72:

GGCAGAGCTT AGAGTGTGGA AAAGGCAACC AGGTGGCCG TAACTGCCTG CTGGAATGCC 60
 50 TGTGCCTCCA CAGGCTCTG GGCATCCGGA CTGATAACCA GTCGGCCAGA CTGAGGGATG 120
 GAAGGCACTG AGATGGBGC CCGTCCAGGC GACACCCGC AGAAATGGAG CTTTCTGTGG 180
 TCTCTTGAC TGTGGTBC CTTTGCCTC TTTGTCTCTC TTTTCTTGG TCTCTCCCT 240
 55 TCTCTCTCTC AGCTTCTCT TCTCTTTGG TGCACCTTA GTTATTGTTG TGAGCAATG 300
 AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GACACAGCC TAAAAAGGA 360
 60 AAAAAAGTAG AAGACAGCAT AGCAACTCAG CTCAGGGAGT TACCAGAGAA AAATAGCAAC 420

	TCATCTTAT GTTCTTCTT TCTTTTAACT TCGATAAAA AGAATTAGAA GTGATCTGCC	480
5	TCATATAAA GCTTCTCTCC CCTTCCCGG TACAGTCTCT TCTCTCTCCC TTAGAGGGA	540
	GAAAGTATG AAGCTTAAAG GATTCTAGC CTGAAAGAG GATCCCGCTG ACCTCTAGC	600
	TATGCAAGAG ACTGAGGCTT GATGCTGGG AGAGGAGAG AAGATCTTG (K)AAGCTTA	660
10	GATATTCTT CAGATTACA GCTCTCTTG GCTAAACAG GTTAGGTAGA CTATCGCTT	720
	TGCGAGTGG CATTCTTGG TACCAATAT TCTGAGGTT TTAGGATTT GCTTGGCTT	780
15	TCTTCTTGG CTTTCTTCT CTTTCTTCT TCTTCTTCT TCTTCTTCT AAGAAAAAT	840
	AAGGCTGCT CCGACTCTG GATGAGGCT CTCACGAGT CAGCATATC GAAGATAAT	900
	TTTAACTGC ATTCTTATG ATTCTTCTT ATCTCTGAT TCTCTCTCT GAGGAGTGA	960
20	GAGGCTCTC AGGAAAGGC ACGACTCTT AGTGAAGTG CTCCCGAGT GATGAGCTG	1020
	GGCATGATG CTGAGGCTG GAGACACAC CTCTGCTCT TCAAGGCTG GCGCTGATG	1080
25	CTCCAGAGT CTCTCTGCT CTAGATGCT CATCTGCCAC CTCTTGTAA GCTCTAGCT	1140
	AGAGGAGAG CCGAGGCTAG AAGAAAGTGA TCTCTGAGA AAAAAAGAT GAAAGTCACT	1200
	TCTACTGAG TCTCTTATA TCTTAAAAA TCTCTATTGA AAGGTAAAA AAGGGGAGG	1260
30	CCCGTACCT AAT	1274

35 (2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 688 base pairs
(E) TYPE: nucleic acid
(C) STRANDEDNESS: double
(I) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

45	GCTACGAGTG GAGGCAATGC CAGCTCCAGG ACAGAGGCTC AGGTGCCCCA CGGCAAGT	60
	AGCCGAGGG GTTCTCTCTG TTAAGGCAG GTTCCCGGG CTTCTGGGCA CAGGCTTCT	120
50	AGGCGAGCT CAGGCCCCA CCGACGGAG TGAAGAGGC GCTTGGGCTG CCATGGTCT	180
	GAGCTTCTC CTGCTGCTGC TCAGCTGGG CAGCTCTGA CAGGCTGCA CAAAAATT	240
	CGAGCTGGG AGAGCATCTA CTGGGGGCT ACAGCTGACA GCGAGACAC ATGCGCTCT	300
55	CTGCTGAAG CAGGCTGCT GCGGCCCCG TCCCGGGCA AGGCTCTGG CCGGAGACC	360
	CTCTCCCGCT CAGCGCGGAC AGCGCCCCG AAGGCGAGAG CTGGAGTGA CGCCTTGGG	420
60	CTTCTACTG TGGCTGGG CTCTCCCG CCGCGGAGG CCGGACCTC TGCACGTG	480

ACCGCGCGCG GGGCGCTCCC TGSTGGCGAT GCGCGGGCAC TGGCCGAGCA CTGCGGGGGG 540
 TTTCCTCTCT GTTGGTTGCT GAGTGGGGGG CCAAGGGGAG AAAAGGAGCC GCTTCTGCCT 600
 5 CCGTGGCCAA AACTCCGTTT CTAATTAAAT TATTTTACTT AAAAAAAAAA AAAAAAAAAA 660
 AAAAAAAAAA AAAAAAAAAA AAAAAA 688

10

(2) INFORMATION FOR SEQ ID NO: 74:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1890 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

20 (ii) SEQUENCE DESCRIPTION: SEQ ID NO: 74:

GAGCAGSAGA GAAGGCACCG CCCCACCCCG CCTCCAAAGC TAACCCCTGG GCTTSAGGGG 60
 AAGAGGCTGA CTGTACGTTT CTCTACTCTT GGCACCACTC TGCAGGCTGC CATGSGGGCCC 120
 25 AGCACCCCTC TCCTCATCTT GTTCCTTTTG TCATGGTGGG GACCCCTTCA AGSACAGCAG 180
 CACCACCTTG TGGASTACAT GSAAGGCGGA CTAGCTGCTT TAGAGGAACG GCTGSCCCAG 240
 30 TGGCAGSACC AGAGTAGTCC GCATCTTCTT GAGGTGCGGG ACTTCAAGAA CAAGATGCTG 300
 CCACCTGCTG AGGTGGCAGA GAAGSAGCGG GAGGCACTCA GAACTGAGGC CGACACCATC 360
 TCCGSSAGAS TGGATCGTCT GSAGTGGGAG GTAGACTATC TGSAGACCCA GAACCCAGCT 420
 35 CTSCCTCTG TAGASTTTGA TSAGAAGSTG ACTGSAGGCC CTGSSACCAA AGSCAAGGGA 480
 AGAAGGAATG AGAAGTACGA TATGTTGACA GATGTGGCT ACACAATCTC TCAAGTGA 540
 40 TCAATGAAGA TTCTGAAGCG ATTGTTGGC CCAGCTGGTC TATGSAACAA GSATCCACTG 600
 GGCACAAACG AGAAGATCTA CGTCTTAGAT GGCACACAGA ATGACACAGC CTTTCTCTTC 660
 CCAAGGCTGC GTGACTTCAC CCTTCCCATG GCTGCCCCGA AAGCTTCCCG AGTCGGGCTG 720
 45 CCGTCCCCCT GGTAGGCAC AGGGCAGTTG GTATATGGTG GTTTTCTTTA TTTTCTTCGG 780
 AGGCTCTCTG GAAGACCTGG TGSAGTGGT GAGATGGAGA AACTTTTCA GTTAATCAAA 840
 50 TTCCACCTG CAAACCGAAC AGTGGTGSAC AGTTCAGTAT TCCAGCAGA GGGGCTGATC 900
 CCCCCCTACG GCTTGACAGC AGACACCTAC ATGACCTGG CAGCTGATGA GSAAGTCTT 960
 TGGGCTCTCT ATGCCACCCG GSAGGATGAC AGGCCTTGT GTCTGGCCAA GTTAGATCCA 1020
 55 CAGACACTGG ACACAGAGCA GCACTGGGAC ACACCATGTC CCAGAGAGAA TGTGAGGCT 1080
 GCTTTTGTCA TCTGTGGGAC CCTTATGTC GTTATAAGA CCGTCTCTGC CAGTGGGCCC 1140
 60 CGCATCCAGT GTCCTTTTGA TGTGAGGGA CCTGACCCC TSAACGGGCA GCACTCCCTT 1200

5
 10
 15
 20
 25

AATTTTCCTGG CAGATATGCT TCTGCTGGCA GGTTCGGCTA TAACCCCGGA GAACGCCAGC 1260
 TCTATGCTG GATGATGAGG TACCATATG TCTATTAAGT GGAGATGAGG AAGTAAGAGG 1320
 AGGAGTTTTS AAGAGCTAGG CTGCTTTTCT GATCTTTCT CACTGCTATA GATTATATAT 1380
 ATATGCTCAT TAAATTCTCT CTCTGCTATT CTTCAAATGT GGGGCACTTG TGCTGCAAAA 1440
 CCTCTATATT TTTAGCCAAT TGTATGAAA TTTTTCAGC TCTTTTCTT CATAGGGAAC 1500
 TCTAGATCTT GAGTAATGCT TTTAGAGGCT GAAGAGTCAA AAGCTGCAAT GTTCCTCTCT 1560
 GTTCTCTCTG CCTATGCTAA GAAATTCAG GTTAAGGATG CCGTAGACCC AGGCTCTCTA 1620
 CTTCTATGCT GGGCAGGGCT AGGAGAGAGG CAGCAGTGTG CTTCTCTCA GATGACTTCT 1680
 GGGAGGGAGA AATAGGAGGA GAGTCTAGT TCTGCTCTCT CTTCTCTACT CTTCTCTCTA 1740
 GTGCTCTAG GAACAGGACT TCTCTCTACT TCTTTCTGAT TCTCTCTCT TGCATTAAAA 1800
 GGAATCTCA CTGCAAAAAA AAAAAAAA AAAAAAAA AAAAAAGG CAGGAGGGGG 1860
 GTTCTCTCTA CCAATGCTCT TCACTGCTAT 1890

30
 (2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1133 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

40
 45
 50
 55
 60

GCGCTCTGA GTGCAGAGCT TCTCTCATGG CCGCTGCTCT GTGCTGCTCT TTTCTCTCT 60
 TCTCTCTCT GCTCTCTCTG GCTCTCTCTG AGAGCTCTGA GTGCTCTCTG GCTCTCTCTG 120
 AAGGATCTGG AGGATCTGG GCTCTCTCTG GAGATCTCTT CAATCTCTG GCTCTCTCTG 180
 TTCTCTCTG GTGCTCTCTG GCTCTCTCTG TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG 240
 AAGAGCTCTG CTTCTCTCTT AAGAGCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG 300
 GATCTCTCTG AGTCTCTCTT GCTCTCTCTG CTTCTCTCTG TCTCTCTCTG GCTCTCTCTG 360
 TCTCTCTCTG AGGCTCTCTG AGGCTCTCTG ATCTCTCTG CTTCTCTCTG TCTCTCTCTG 420
 TCTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG 480
 GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG 540
 CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG CTTCTCTCTG 600
 GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG 660

221

AGTTCATGAC AAGACTCTTC TCTCAAAAT CATCTGGCAA ATCTAGCAGT GGCAGCAGTA 720
 AACACGGCAA AAGTGGGCT GGCAAAAGGA GGTAGTCAGG CGTCCAGAG CTGGCATTTG 780
 5 CACAAACACG GCAACACTGG GTGGCATCCA AGTCTTGGAA AACCTCTGA AGCAACTACT 840
 ATAAACTTGA GTCATCTGGA CGTTGANTTC TTACAACTGT CTATTTTAAC TTTTTCAGAC 900
 ATGTTTGTGA CTGGTACAC GAGAAAATCC AGCTTTCATC TTTTCTCTGT ATGAGGTCAA 960
 10 TATTGATGTC ACTGAATTAA TTACAGTCTC CTATAGAAAA TGCCATTAAT AAATTATATC 1020
 AACTACTATA CATTATGTAT ATTAAATAAA ACATCTTAAT CCAGAAAAAA AAAAAAARAA 1080
 15 AACTCGAGGG GGGGCCCGGT ACCCAATTTC CCAATGGGA GTCGTAAAAA ATC 1135

20 (2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 585 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76

30 ATGTTTACAA TGTGTGTAT AAATGGGACA ACTCCTCGCC CTCTACCTGT CCCCTCCCCC 60
 TTTGGTTGTA TGATTTCTCT CTTTCTTAAG AACCCCTGGA AGCAGGCGCT CCTTCAGGGT 120
 TGGCTGGGAG CTGGGCCCAT CCACCTCTTG GGGTACCTGC CTCTCTCTCT CCTGTGGTGT 180
 35 CCCTTCCCTC TCCCATGTGC TCGGTCTTCA GTGGTGATA TTTCTTCTCC CAGACATGGG 240
 GCACACGCCC CAAGGGACAT GATCCTCTCC TTAGTCTTAG CTCATGEGGC TCTTTATAAG 300
 40 GAGTTGGGGG GTAGAGGAG GAAATGGGAA CCGAGCTGAA GCAGAGGCTG AGTTAGGGGG 360
 CTAGAGGACA GTGCTCTCGG CCACCCAGCC TGTGCTGAGA ACCATTCTTG GGATTAGAGC 420
 TGCCTTTCCC AGGGAAAAAG TGTCTCTCC CCGACCTCC CGTGGGCCCT GTGGTGTGAT 480
 45 GCTGTGTCTG TATATCTAT ACAAAGGTAC TTGTCTTTC CTTTGTAAA CTACATTGTA 540
 CATGGATTAA ACCAGTATAA ACAGTTAAAA AAAAAAAAAA AAAA 585

50

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 577 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 60 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 75:

5 GGCACGAGGC CTGACGAGAG TTCTACTTGG CTGCTTCCCT GCTCTTGGCC ATGAGCTTGGC 60
 GTTGGCTCAG CTGCTTCTTG ATGAGGAGCT TCTTCTAGT TTCCGAGAGA GTCTTGGGCT 120
 AGCTGATGGT ACTGCTGGCC TTCCAGGGG AGGTGCTGA ATTCTCTTGT AGGTCAGG 180
 10 CCGAGTAACT CAGCATTAGA GATTACGCTG TGCTTTGTA CAGGAGGGG GCAGGCGATA 240
 CCGCTGATA TTCTCTTAT TATCGCTGG AGGAGGATGA CCACTGGGCT GCTGACATCT 300
 CCGATCGATT CTGACAGCT AAGGATGAG CCGACATGG CTGTCTCTTC ACCATTAGTA 360
 15 CCGTGCAGCT TGAAGACGAT GCGATTACT ACTGCTTCT TGGTATAGC TTATAGTCCC 420
 AGAGTGGGG TCGAGATGG GTGCTTCCC TTGCTCTCT ACTATGCTC CTGACCTTGC 480
 GCTCTTTTA AACTTCTCT GAGCTTGGT TCGCTTCT AAGAGGCTT AATAATTT 540
 20 AACCTGTAAA CAGCAAAA NAAAAAATA AACTGG 570

25

(2) INFORMATION FOR SEQ ID NO: 76:

(1) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 2278 base pairs
 (E) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 78:

35

GTAATTGGG AGGATGGGC CACATGGGG GGTGGGGCT GGGGCGGCA CTAACGGCC 60
 CTCTTGAGG CCGATAGGG GGTGTGGG GCGAGGAG GCGCGAGA GCGCGGCT 120
 40 CCGTGGAGG AGAGGGGT CCGCTCCAT ACCGCTCCA ACTGAGGCT GTGATGGAG 180
 GGTGAGTGA TGCTAAAT TTACGCCCC TGTCTGCT CTTCCACCA GACTGATTCA 240
 GAATGGGAG CTTTCCAAA GAATGCTGA ATACTTGA CAGTGTGG GAAGGTAGAC 300
 45 GTGATTAAG AATGAGTTC GAGTGGGG TTCTTTTCA CACTCTCC AGCATTTT 360
 CATGCAAGG ATGAGATAT CCGCTTAT GGTGGGAG GAATTTGA AGACTTGCAG 420
 50 AATCATATC TAGAGAGAA ATGCAATCA GTGAGGTC GAGTGGTG GAAATCCCG 480
 GCTTCTTAA CAGTCTTGG AATGCTGCT CTTTATGCA CTTCTGCAA GATATAGCAT 540
 CTTCAAACT ATTTACAGT GACTTTGGA ATTCTGCT GTGTCTTA TGTCTTTT 600
 55 GTCATAGCA CTTGCTTCT TGGCTTTT ATGCTGCG GTTGGTGT AATATCAGAG 660
 TGTTCCTTG TCGACTTCC AAGCATTT TCTAGGCT CTGAGCAGAA TGGAGATCA 720
 60 GAGGAGGCT ATAGAGCTGA ACAGTTCAG GATCGGAGG AGGAAAAA TATTCAAAT 780

GAAGAAGAAA AATAAAGACAG CCTCTTAGAT GATBAAGAAG AGAAAGAAGA TCTTGGGAG 840
 GAGGATGAAG CAGAGGAAGA AGAGGAGGAG GACAACCTTG CTGCTGCTGT GAGTGAGGAG 900
 5 AGAAGTGAGG CCAATGATCA GGGGCCCCCA GGAGAGGAGG GTGTGACCCG GGAGGNAAGT 960
 AGAGCCTGAG GAGGCTGAAG AAGGATCTTC TGAGCAACCT TGGCCAGCTG ACACAGAGGT 1020
 10 GTTGAAGAGC TCTTGAGGC AGGTAAAAG TCAGCATGCT GNDAGGGAC TGTAGATTTA 1080
 ATGATGCGTT TTTAAGAATA CACACCAAAA CAATATGTCA GTTCCCTTTT GGCTGCAGT 1140
 TTTACCAAAA TTTTAAATT TCCCTGAATG AGCAAGCTTC TTTAAAAGA TGCTCTCTAG 1200
 15 TCATTGCTG TCATGGGAGT AAGGCTCATG TATACTAAGG AGAGCTCTCC AGGTGTGACA 1260
 ATCAGGATAT AGAAAAACAA AGCTACTGTN TGGAATCTGT TTGAGACTG GGATGGGAAT 1320
 20 AAGTTCATTT ACTTAGGGGT CAGAGAGTCT CGACCAGAGG AATCCATTCC CAGTCCTAAT 1380
 CAGCACCTTC CAGAGACAAG GGTGCAAGGC CTGTGAAATG AAAGCCAAGC AGGAGCCTTC 1440
 GNTGTGAGGC ATCCCAAG TGTAACTAG AAGCCTTGA TCTTTTCTT GTGTAAAGTA 1500
 25 TTTATTTTTS TCAAAATGCA GGAACATCA GGCACCAGAG TGATGAAAA ATCTTTCACA 1560
 GCTAGAAATT GAAAGGGGCT TGGGTATAGA GAGCAGCTCA GAATCATCC CAGCCCTCTT 1620
 30 AATCTCTGT GTTATCTTT ATTTCTIACC TTTAATTTT CAGCATTTT CACCATGGG 1680
 ATTCAGGCTC TCACACTCT TCACTATTAT CTCTTGGTCA GAGGACTCCA ATAACAGCCA 1740
 GGTTTACATG AACTGTCTTT GTTCATTCTG ACCTAAGGGG TTTAGATAAT CAGTAACCA 1800
 35 AACCCCTGAA GCTGTGACTG CCAACATCT CAAATGAAAT GTTGTGGCCA TCAGAGACTC 1860
 AAAAGGAAGT AAGGATTTTA CAAGACAGAT TAAAAAATA TTGTTTTGTC CAAATATAG 1920
 40 TTGTGTGTA TTTTTTTTA AGTTTTCTAA GCAATATTTT TCAAGCCAGA AGTCCTCTAA 1980
 GTCTTGCCAG TACAAGTAG TCTGTGAAG AAAAGTTGAA TACTGTTTTG TTTTCATCTC 2040
 AAGGGGTTC CTGGCTCTG AACTACTTTA ATAATAACTA AAAAACCCT TCTGATTTTC 2100
 45 CTTCAGTGAT GTCTCTTGG TGAAAGAATT AATGAACTCC ACTACCTGAA AGTGAAAGAT 2160
 TTGATTTGT TTCCATCTTC TGTAATCTTC CAAAGAATTA TATCTTCTA AATCTCTCAA 2220
 50 TACTCAATCT ACTTAAGTA CCCAGGGAGG CTAATTTCTT TAAAAAATA AAAAAAA 2280

55 (2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1143 base pair:

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

60

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

5 CCGCTGCAAG TGTAAACCA GTTCTCAGG CAGACGCGA GCGTTCGGG CCGCGGTTCG 60
 CAGGTTCGGA GAAATGAGA GTGTCTGCT CACAAATGGA ACACGTGACA GTGATGAGAT 120
 GGGCTGGTAG GTCCCAACG ACATCCACAT AGATGTACG ACCAACAGG ATTCTGCTCT 180
 10 GGCAAGATC CTGTTCACCT GTTCTCTGG GCTAGGCGG GGGGCGACCC AAGCAAGGG 240
 CAGAACGGG CTGTGTGTG CTTCTGGGT CATCAACAG GCGCTGGGA TTTTGAATG 300
 15 AGTCTAATA GAATTTTCTT ACATCCGCA CTACACCTC CIGCTACCT CAGGAGCTG 360
 CATCTGATA GGGGTGTGG CTGTCTGCT CGAGCTGCT GCTTCATTT AGAATAAG 420
 GGGTCTACA TACTGGGCGT TACTGAGGAT TCTCTAGCG CTGGCAGCTT TCTCTACAG 480
 20 CATCTCTAG CTAATAATT GAAATGAGA TTTTGATAT GGTACTCTT ATTAACAGC 540
 TGGCTGCGG ACCTCAGCT CAGCTGACG GAACTCTCA GCGCCCACTC AAGTCCAG 600
 25 AGAATTCAG AGATACACC TATGTACCT GTTCTGGA ATGCTGAAG CTTTCTTCA 660
 AACCTTCTG GCGATGCTT TGGGTCTCT GATCTCTCT CTTCGGCAT CTCTGGCCC 720
 TCTGCGCTG TACTGCTGA GATGCTTCC AACCAAGGG AAAAGAGACC AAGAGGAA 780
 30 GTTGAATG AGAGGATCT AGCATGCT CTCTGATTA TTAGTGCCTG GTGCTTCTG 840
 ACCGAGGCTC CTCTCATCTG ACTGCTGAA GAAGAACCAG ACTGAGGAAA AAGGCTCT 900
 35 CAACAGCTC AATATCTG GCGCATGAC CCGGCGACA GCGCTGCTC ACAGGACT 960
 GCGCATCTT TAAACCTT CCGCATCTG CTGCGCTCA TCTCCCTCC TAGTATCTA 1020
 TGTGAAATA AATCTCATG TTATCTTNN NAAAAAAAAA AAAAAAAAAA AATTTGGG 1080
 40 GGGCTGCTA CCAATGGG CTGCGGCGG GCTTAAAAA TAATGGGGG GCTTAAAAA 1140
 GGT 1147

45

(2) INFORMATION FOR SEQ ID NO: 80

50

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 517 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

GGCAGACAGC AGATGGCCTT GACACCAGCA GGGTACATC CGGTATTGCT ACTTCTCTG 60
 60 TCCCCACAG TTCTCTGGA CTTCTCTGGA CCACAGTCT CTGCAAGACC CCGCCAGAG 120

225

5 CCGAGTCCAC CATGATCCAT CTGGGTCACA TCTCTTCCT GCTTTTGCTT CCACTGGCTT 180
 CAGCTCAGAC GACTCCAGGA GAGAGATCAT CATTCCTGC CTTTACCTT GGCACCTCA 240
 GGTCTTGTTC CCGATGTGGG TCGTCTCTCT TATCGCTCCT GGTAGGCTT GTGGCTGCTT 300
 ATGCGGAGG ATGCTGTTC ATGGGGAAG GGTCTTCCT GTGTCACCT CCAAGGAGC 360
 10 GTTCGAGGCA AGAGATGGT AAATTTTADA TCAACATGCC AGGCAAGGCT TGAACCTCCT 420
 GAGCTTGA CTTTGAATT CTGACCTCT CATCTGGAT GGTGTGTGT GGCACAGGA 480
 CCCCCGCCCC AACTTTTGA TTTAATAAA ACAATGAAA CACCAAAAAA AAAAAAAAAA 540
 15 AAAAAAAAAA AATCGA 557

20

(2) INFORMATION FOR SEQ ID NO: 81:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 795 base pairs
 (E) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (ii) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

GCGGAGAGGA TGTGAGCGC GGGCGGGG GGGCTGCCT GGGCGTGT GTTGGGGCTG 60
 CTGTGGGGG TTTAGTGGC GGGGGTGT GTGGCAAGA CCGGTGGGA CTGTGACCT 120
 35 GCGGTGGGT GCTGAAGCTG CTCAATAGC ACCACCGCT GTGTGCACT GGCACGACAT 180
 CAATAGGGA TGGGCAGCG GCGAGCATC GGTGACCGG GTAGAGGGT GGCACAGCT 240
 MAATAGGTAC TGGGGATCC GGGCGGGT GAGGGCGGG TGCCCGCGG GTTCCCGGT 300
 40 GTGTGGGG CAGGGGTGA GGTCAAGCA TTTCTTACG GCAAGAACY TGCACAGCA 360
 CCATTTCCG TGGCGGTGT CCAACAACA GAGGTGAGT GCTTTGGGG AAGACGGGA 420
 45 GGGGAGGAC CTGACCTAT GACAGTGG CTGTCTGGA CAGCACTGG AGCTGAGGT 480
 TGTGTGTGT TCCAGCATGT GGCACCTCT GTTCTCTGT CAGTCAAGG TGACAGTAT 540
 GGAAGGCGCA TGGTGGGCA GCATGAGGT CAGGCTATG CAGTGGCA CATTACAT 600
 50 AGGTGGAAG CCATGGAAG CATTTTATC AAGCCTACT TGGAGCCTT TGCAGTCA 660
 GATGAATCT GATGTGTGG ATGATGGGT GATGAGGG TGGAGGTG GGTGTGTGA 720
 55 GGGCACTCT TGCAGAGAC TTGGGTCTG TAGGGTCTT CAATGCTT TTGATTAAA 780
 GAATGTGT CTATG 795

60

(1) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1324 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 82

5 NAGGCTTAA AGCGCTTACC CTNCTGCAG GTGAGCAGTG GTGCTGAGA GCCAGGCGT 60
 CTCTGCGTG CCCACTCAGT GGAACACCC GGSAGCTGT TTCTCTTTG TGSAGGCTCA 120
 15 GCACTTCCT CTTCAGAAC TCACTGTCAA GAGTCTGAA CAGGAGCCAC CATGCACTG 180
 TTAGATTAA TTAAGACCAI GAGGATCTC TTAATTTCG TCACTCTTCI GTGTGGTGCA 240
 20 GCTCTTTTG CAGCTGCAAT CTGCTGTCA ATGATGGGG CATCTTTCT GAAGATCTT 300
 GGGCACTAT CCGCAGTGC CATGAGTTT GCAACGTG GATCTTCT CATCGCAGC 360
 GGGTTGTG TCTTTGCTCT TGGTTTCTG GGTGCTATG GTCTAAGAC TGAAGCAAG 420
 25 TGTGGGCTG TGACTTCTI CTGATCTC CTTCTATCT TATGCTGA GGTGCGAGT 480
 GTGTGTTG CTTGTGTA CACACAATG GCTGAGCACT TCTTACCTT GCTGTAGTG 540
 30 CTTGCTACA AGAAGATTA TGGTTCAC GAAGACTCA CTCAGTTTG GAACACNAC 600
 ATGAAGGCG TCAAGTGTG TGGTTTACC AACTATACG ATTGTGAGA CTCACCTAT 660
 TCAAGAGA ACAGTGTCT TCTTCTATC TGTGCAATG ACAAGTCA CACACAGCC 720
 35 AATGAACCT GCAACAACA AAAGTTTCA GACCAAAAAG TAGAGGTTG CTTCAATCA 780
 CTTTGTATG ACATGGAAC TAACTAGTC ACCGTGGGT GCTTCTAGC TGGAACTGG 840
 40 GGGCTGAGC TGGCTGCCAT GATTCTKTC ATGTATCTGT ACTGCAATCT ACAATAAGT 900
 CATTCTGTC TTGTCCTA CTGCTGCCAC ATGGGAATG TGAAGAGCA CCGTGGCAAG 960
 CAGCACTGAT TGGGAGAGG GACAGATCT AACAAATGCA CTTGGGCTAG AATGACCTG 1020
 45 CCGTTCTGC TCCAGACTG GGGCIAGAT GGGACCACT CTTTAGGGA TGGTGAAT 1080
 TCTTCTATT GATGCTGA TGGTCTGAG GATTTCCAG GGTCTAAG TATCTGTT 1140
 50 TGTGCTCAT TGGGCTAGT TATTAACCC TTGATATGC CCGTAGGCT ATGCTGATC 1200
 CCACTCTCT ACTGGGGAT GAGAGAAAG CATTTATAG CCGGGGATA AGTGAATCA 1260
 55 GAGAGCTC TGGTGGAG TGTAAAGG ACTTCAAAAT GATTAACCT GTTACATCT 1320
 TAAA 1324

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1494 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 83:

10 CTCAGGCTTC TGTCTCACTT TTGGGGGGGG GGGATTAGGG CAAGSAGGGC ATGAGGGACT 60
 GTCTCTCCCT AAAACCCAGA CCGGTGTGCC CCACTCAGTT CTTCTTCATC CTCTCTCTCA 120
 15 TCTTCATTGC TGAGGTTGCA GTTGTCTGTG TGGCTTGGT GTACACCACA ATGCTGAGAC 180
 ACTGCGATGS AGGAAGGGA GAAGATTGGG CAAAACCCCTG GGAGTGGGCT GTGGCTGTG 240
 AATGCCACCC TTCTGTACCA GGGCTAAAC ACTGGCTGTC CTCACCCAGG CTGAGCACTT 300
 20 CCGGACGTTG CTGCTAGTGT CTGCGATCAA GAAAGATTAT GGTTCGAGG AAGACTTCA 360
 TCAAGTGTGS AACACCAACA TGAAAGGGGT AAGGTTGGCT GGGGGAGGTT TTAGGCTGGA 420
 25 GAGAAAGAAG CAAGGCCCCA CTTCCACCTT CATCTGTCT CCAGTCAAG TGTGTGGCT 480
 TCACCAACTA TACGGATTTT GATGACTAC CTTACTTCAA AGAGACAGT GCGTTTCCCC 540
 CATTTCTTTS CAATGACAA GTCACCAAC ACAGCCCAAT GAAAGCTGCA CCAAGCAAAA 600
 30 GGCTCACGAC CNAAAARTAN AGGTGTGGG TGGCATGAGT GGGTGGGAC TGTTTTCATG 660
 GCTCAGAGT GGCAACGGG GATGGAGTA GGGCAGGTGC CAATTATAAA TGTCTTTTTC 720
 35 TCTTCCYGAA GGGTGTCTT AATCAGCTTT TGTATGACAT CCGAATAAT GCAGTCACCG 780
 TGGGTGTGT GGCAGCTGGA ATTGGGGGG TCGAGGTAAG CAGATGAGG GCTGGGACTG 840
 GSACATGGGC ATGAGACCAG GGTGTCTCA CCCATCTGAG GCTCTCTTG AGGAAACAGA 900
 40 CTCTTAAGTG GGCTCAGET AGGTGTCTG TGGGACAGGC TTCAGGATCC CTATCATGTT 960
 CCGTCATCTC TCGGTGTTC TGGCTCTCA GCTGGCTGCC ATGATTGTGT CCATGTATCT 1020
 45 GTACTGCAAT CTACAATAAG TCGACTTCTG CTTCTGCCAC TACTGCTGCC ACATGGGAAC 1080
 TGTGAAGAGG CACCGTGGCA AGCAGCAGTG ATTGGGGGAG GGCACAGSAT CTAACAATGT 1140
 CACTTGGGCC AGAATGGAAC TGGCTTTTCT GTTCAGACT TGGGGTAGA TAGGGACCAC 1200
 50 TCGTTTTAGC GATGCCGTGAC TTTCCTTCCA TTGGTGGGTG GATGGGTGG GGGCATTCCA 1260
 GAGCTCTAA GGTAGCCAGT TGTGTGCCC ATTCCCCAG TGTATTAAC CTTGATATG 1320
 55 CCGCTAGGC CTAGTGTGA TCGAGTGT CTACTGGGG ATGAGAGAAA GGCATTTTAT 1380
 AGCTGGGCA TAAGTGAAT CAGCAGAGCC TCTGGGTGGA TGTGTAGAAG GCACTTCAAA 1440
 60 ATGDATAAC CTGTTACAAT GTAAAAAAA AAAAAAAA AACTCGACTC TGGC 1494

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1285 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 84

5	GCTACGTGGG TGGACGGAT GGAAGGAGG CCGTGGGGG GAGTTGCTT CTGCTCCGGA	6
15	TGCAATTCTT GTGGATGAG TTCTGGGAG GGAAGGAGG GGTGAGGCG CTGCTCTCTT	12
	AGATGGGAT CTAGCTGAG CCGTCCATGA ACGGTGAGG CTAGGAGAT GCGTACCA	18
20	GGGGTTCAGA TGGGTGGG TGGGGGAGG GCGGTGGAA CAAGGAGAG ATGATCTTA	24
	ACCAATATCT TACTGATCT AACACAGGAC TTGGGGAAG ACAGGAGAT GGGAGGTTG	30
	CCGACATGCT CCGCAAGCAT CAAGTGGCAT TGCGCACTTA CTACACCGTG CCGAATGGCA	36
25	CCGTGGCTCT TGAAAGGAG GCACTAATCA AGTGGATGAA GGGGATCCCC TTTGTGCTAA	42
	GTGGCAAGCT CAGGAGGAGT GAGCTGGGG TTCTCTACCC ATTGACATG ACTGGCACC	48
30	CGTGGGCTGG CCGTGAGTTC AGCGCCACAC CAGATGATGC TTCTTTTGG TGGCTCAGCA	54
	CTGTCTATGC TGGAGTAA CTGGGATGC AGGACACCAG GGGGAGCCC TGCCACAGC	60
	AGGACTTCTT CAGGAGGAG AACATCATCA AGGGGGCTG ACTGAGACA GGTCTCCCG	66
35	GAGGATGAA TGAATTCAGT TACCTACACA CCAACTGCTT TGAATCACT GTGGAGCTG	72
	CTGTGACAA GTTCTTAC GAGATGAA TTGGGAGGA GTGGAGAAC AACAAAGAG	78
40	CCCTCTCTAC CTAGCTGAG CAGGTGGCA TGGGCTATTG AGGAGTGGTG AGGACAAAG	84
	ACAGGAGCT TGGATGCT CAGGTGCA CTGCTGGGA TGGATTAA CAGGAGTGA	90
	CCAGGAGCT GGGGGGAA TATGGGCTT TGTGACCC AGGGGACTAC ATGTTGACT	96
45	CCAGTCCGA GGGTACAT TCACTGAC GGAATGCTG GTTCACTTT GAAGAGGGCT	102
	CTTCTCTCT CAACTTCTT CTCAAGA CTCTAAAGG GAGGTGGG GAGTCTCTG	108
50	CAGCTGGGG CAACTTCTT CCGGAGCTG GAGGGGCTT GAGGGGCTA AGGGGAGG	114
	AGGATTGATA CTTGAGGTTT AAGAGCTTA GGGGAGCTT GAGCTCTAA GAGGGAGG	120
	GGAAGAGTAG AGGGGAGG ACAAAGTGA GAAAGGCTT TCTTAAAG TACCGGGAC	126
55	CTTAAAAAA AAAAAAAN AAAAA	128

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 394 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 85:

10 GCGCGCTCTA GGAAGTACTG GATCCCCCGG GNTGTCAGGT GTGGAGTGGG CCATCGTAAA 60
 TAGTATCTGT GCATAAGGTG GTTGTGCSAT AAATGAGTTA ATATATGCAA AGCCCTTGGG 120
 15 CCAGAGCCGG CGCAGAGGAT TGTSTAAGTS CTGGGAGGGG TCATGATGGA GATATCATGT 180
 CTCCTCTTCT TGAATCAGGA TTCTGATGAG ATGGAGGATG GCCTTGGGTT TCAGGATTAG 240
 GCCTTGAGGC ACTGCTCCAG CTTCTTTTGT GGGCCCTGTC ACCCTTGGCT TCATCGGGCC 300
 20 GTARCAAGTC TCCCCCTCCC CACTYTGAG CAGAGTGTT CAAGAAGTGC CTGCTCAGCG 360
 TTCGTGTCTT GCAAGGCCAT CGCTAACCT CTAA 394

25

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1925 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 86:

30 AGTGAAGGGA GCTGGGCTTG CCACTGGGCT TCGGGCCCTG TCCAGAGGA GCANGCCTTC 60
 40 CTGAGCAGGA GGAAGTAGST GTTGCCCGG GCCTTGAGGC AGCCCTGCA GCTGGATGGA 120
 GACCTCCAGG AGSATAGAT CCCAGTGGTA GCTATTATGG CCACTGGTGG TGGGATCCG 180
 45 GCAATGACTT CCTGTATGG GAGCTGGCT GGCCTGAAGG AGGTGGGCT CTTGGATTG 240
 KTCTCCTACA TCACCGGCGC CTGGGGCTCC AACTGGGCT TGGCCAACCT TTATAAGGAC 300
 CCAGAGTGGT CTCAGAAAGA CTTGGCAGGG CCACTGAGT TCTGAGAGC CCAGGTGAC 360
 50 AAGAACAAGC TGGGTCTGCT GCGCCCGAGC CAGCTGCAGC GSTACCGCA GGAGCTGGCC 420
 GAGCCTCCCC GCTTGCGTA CCAAGCTGC TTCACCAACC TGTGGGCTT CATCAACGAG 480
 GCGCTGCTGC ATGATGAGGC CCATGATCAC AAGTCTCAG ATCAACGGGA GGCCTGAGT 540
 55 CATGGCCAGA ACCCTCTGTC CATCTACTGT GCGTCAACA CCAAGGCA GAGCTGAGT 600
 ACTTTTGAAT TTGGGAGIG GTGGAGTTC TCTTCTTAG AGGTGGGTT CCCCAGTAG 660
 60 GGGGCTTCA TCCCCCTGGA GCTCTTGGC TCCGAGTCTT TTATGGGCA GCTGATGAAG 720

5
 10
 15
 20
 25
 30
 35
 40
 45

AAGCTTCTTA AGGCTGKAT CTGCTTCTTA GAGGATATCT GAGGCAACT CTATGCAAGT 780
 AATCTGAGG AGGATATATA CTGGACCTCA GAGCTTASG ACTTCTGASA TCGTNGGT 840
 AGGACACAG CCAACTTGA CAGGAGGAG GTCCCTCTT TGAAGATAGA AGGACAGG 900
 TGAAGAGAT GAGATATAG TGAGTTTCTT AGGAGATCTT TGAAGTGAGG TCACTTGGT 960
 CAGGACAGC AGATATATCT GTGTGGGTC GATTGACACA AAGATATCT TGAAGATCT 1020
 CATTATATA CATGAAAA TCAGACTCTT GATGATCTT CCAATCAACT GAGACCTTA 1080
 GAGGCTCAGT TCGGCTCTT GAGTCTGCTT TATCTATCA ATAGAGGTG CTGCTCTT 1140
 CTGCAATCA TCAAGATCT GAGATATCT CTGATCTT ACTAGATCT CCAATGAGT 1200
 TCTCAAGT TGAAGTCTT GAGGCTCTT TCAAGAGT AGGATATCT GTTCTCAGT 1260
 AATCTGAGC AGGCTGAGG GAGTCTTCTT GATGATCTT GAGGATCTT CTGCTCAGT 1320
 AATCTGAGG GAGGCTCTT GAGGCTCTT GATGATCTT GAGGATCTT CTGCTCAGT 1380
 TATCTGAGT CTGCTCTT GAGGATCTT GAGGATCTT GAGGATCTT CTGCTCAGT 1440
 TCTTATCTT ACTCTCTTA CCACTACAG AAGTCTCTT ACAGCTCAGG GAGCTCTT 1500
 AAGCTCTT ACTCTACAG TCAATCTT TCAATCTT AAGCTCTT GCTCTCTT 1560
 CTGCTCTT GAGGATCTT GAGGATCTT GAGGATCTT GAGGATCTT CTGCTCAGT 1620
 CCACTCTTA TCACTCTTA TCTCTCTT GCTCTCTT AAGCTCTT TCTCTCTT 1680
 CAGTCTCTT GAGGATCTT GAGGATCTT GAGGATCTT GAGGATCTT CTGCTCAGT 1740
 AGCTCTCTT CCACTCTTA TCTCTCTT GAGGATCTT GAGGATCTT CTGCTCAGT 1800
 CAGGATCTT CCACTCTTA TCTCTCTT GAGGATCTT GAGGATCTT CTGCTCAGT 1860
 CAGTCTCTT ACTCTCTT CCACTCTTA TCTCTCTT GAGGATCTT GAGGATCTT 1920
 CTCTA 1920

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1818 base pair
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 87:

CCGGGCCCC CCGGCGATTT TTTTCTTTT TTTTCTTTT TACGAGTCTG TATGATCTA 80
 AGTCTCTCA CTACTCAAG TAGCGCAGG GGGGAAACAG GCACAGGCCG GGGGCTTTT 120

60

	GGTGATTADA CAAATBGGGT TGAGCTCTCTT ACCCCACTGC AAAGTCTCTGA GGGGCAAGG	180
	AGCTCCACAC CCGCAGGCTG GACCTCAGCA CAGTCCCAAC TGAGTCGAG GCTCTCNAAC	240
5	CACTGCTTCA TCACATCTCC CACCTCAAC TCAGCAGCCA GCGATGAAT GAGAGTTAG	300
	GGTGGAGGG GCGTCGCTCA TCACTCAGGC CTGCGCTGC GCTCTCTCTG CACTGCTCTG	360
10	CGTAGAGCTT CAGCAGCTG CTCCTCTCTG TCGCAGACAT TCTAATGTCT CAGCTCTAG	420
	AGTTTCTTCA CTTCTCTCTG GCTCTCTCTG AGCTCTCTCT AGTCTCTCTG AGAGTCTCTG	480
	GAAGACAGCT TCACCTCTCC AGCTCTCTCT TCTCTCTCTG TCGCTCTCTG CCGCTCTCTG	540
15	GAGTCTCTCC GCAAGGAGTC GCTCTCTCTG GCAAGGAGTC GCTCTCTCTG GCTCTCTCTG	600
	CAGCTCTCTG CCGCAGGCTG GTCTCTCTCT CAGCTCTCTG GCTCTCTCTG GCTCTCTCTG	660
20	GAGTCTCTCT GCAAGGAGTC CCGCTCTCTG TCGCAGACAT TCTAATGTCT CAGCTCTAG	720
	TCTAAGCTCT CTTCTCTCTG GTCTCTCTCT GAGCTCTCTG GCTCTCTCTG GCTCTCTCTG	780
	TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	840
25	GGTCTCTCTG GTCTCTCTCT GTCTCTCTCT CAGCTCTCTG TCTCTCTCTG CAGCTCTCTG	900
	AACTCTCTCT GTCTCTCTCT GTCTCTCTCT AAAGTCTCTG CAGCTCTCTG CAGCTCTCTG	960
30	GCTCTCTCTG GTCTCTCTCT GTCTCTCTCT TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1020
	GCTCTCTCTG GTCTCTCTCT GTCTCTCTCT TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1080
	TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1140
35	AGCTCTCTCT CAGCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1200
	AGCTCTCTCT TCTCTCTCTG CAGCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1260
40	GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1320
	AAAGTCTCTG GCTCTCTCTG TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1380
	CCATCTCTCT GCTCTCTCTG TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1440
45	AGCTCTCTCT TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1500
	CGCTCTCTCT GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1560
50	TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1620
	TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1680
	GAGCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1740
55	TCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1800
	GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG GCTCTCTCTG	1860
60		

(1) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 539 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 88:

AGGGTAATTA ATATGAAGTG CAAAAAGTTT AATGTTCCAG TTTAAAAGGC AGTGGGAGAA 60
 ACTACATAGC ATGGAATAAA TAAATGAAAT TTTTATTAAAT GAGAAAGAGT YTGTTCAGT 120
 GGGAAATGCT GCTGCTCAGC CGATGGGAAI GGGAGCCTTT CAAGCTTTT TTTGGTAAAT 180
 ATTCACAGTT TTTAAGCTCT GTTACTTTT CAAAACGAGC TGTCTCTTCC TTCTGACACC 240
 GATTTGAAAG CTCATGCTG AGGAGAGGT GTTGTGAAGG TCACAGGTTT CTGCTTGCA 300
 TTGTCATAGC GTCTGTAGT ATCACTTGT AGGCACTGTC TGGTTGAAGG AACTAAGAGC 360
 ATTGAGGGAT AGAGAGCTGA AATAGGAT TATTNNTTCC TTTGACTCT CCGCTCAAGA 420
 TGTCTTGT TTTGCTGAA AAGCTCTCT GACAACITTT GTGCAAGCA AACCATCTGC 480
 GTTCTCTGAA CTCTGACTGA ATATATTAGG ATTTCCCTT CTGAGGCTTC GTACTGCCA 539

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 555 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 89:

CTCTGCCCCA GGTGGCAGCC GAGTTCAGGC TGGTGCCAC CACCAAGTT CAGTGGCGT 60
 ACCAGTGGCT TATGCTGCT CCTCAGGAG GGTGCGACA GGNACTTGA CTGCAGCGAT 120
 GGTAGGAGTG AGGAGGAGTG CAGGATGAG CATTGTACCC AGAAAGGCA ATGCCCCAGG 180
 GGGCTGCTC TCCCTGCTC CTGAGGAGG GTCACTGACT GTCTGAGG AACTGATAAG 240
 AAAGTGGCA ACTGAGGCT CCGGCTGCT CTAGGAGGG AGTCTGCTT CAGGTTAGT 300
 GATGACTGCA TTCACTTAC GTGCTCTGT GAGGCTCAC CAGACTCTC CAGCTCCAGT 360
 GAGGAGCTG GCTGTGAAAC CAATGAGAT CTCCCGGAAG GGTATGCCAC AACCATGGGG 420
 CCCCCTGTA CCGTGGAGAG TGTACCTCT CTCAGGAATG CACCAACCAT GGGGCCCCCT 480
 GTGAACCTG GAGAGTGTCT CCTCTGTGG GAATGCCACA TCTCTCTTG CCGGAGACCA 540

233

GTTGTGGAAGC CCAATTGGCT ATGGGGTTAT TNCAGGTGCT GGGGTGCTCA GTGCAAGGCT 600
 GGTGATCGGG ACCCTCTCTCT TTTTGTCTTG GTTGGAGGC CAGGAGCGCT TCCGCTCCT 660
 5 GGGTTTACCT GTGGCCATGA AGGATGCTCT GTGTGTGTCA GAACAGAAAG CCTCGCTGCT 720
 CTGAGGATAA GCACTTGCCA CCACTGTGAC TTAGGCTGCT GGTATGGA CAGGAGGAGA 780
 GGAATGATGG GGATGGGTAC CCAATGAGAG AGGCTTTCAG AGACTTGAGC NCTTCTGGCT 840
 10 ACTGGAACCT CGAAG 851

15

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 628 base pairs
 20 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 90:

AAGGACGTGC CCGCCGCTG GCTTCTGAGC CGGAGTGGTC GGTGGGTGGG ATGGAGGCGA 60
 CCTTGGAGCA GCACTTGCAA GACACAATGA AGAATCCCTC CATTTTGGGA GTCTGTGCA 120
 30 CAGATTCACA AGGACTTAAT CTGGGTGGC GGGGACCTT GTCAGATGAG CATGCTGGAG 180
 TGATATCTBT TCTAGCCGAG CAAGGAGGTA AGCTAACCTC TGACCCCACT GATATTCCTG 240
 TGGTGTGTCT AGAATCAGAT AATGGGAACA TTATGATCCA GAAACAGSAT GGCATCAGG 300
 35 TGGCAGTGCA CAAATGGCC TCTTGTGCT CATATCTGTT CTTCAGCAGC CTGTCTAGG 360
 AACTGGATCC TACCTATBTI AATTACCTTA TAGAACTACT AAAGTTCCAG TAGTTAGGCC 420
 40 ATTCAATTAA TGTGCATTAG GCACTTTTCT GTTTATTTAA GAGTCAATTG CTTTCTAATG 480
 CTCTATGAC CCACTATCAA GATATTASTA AGAAAGGATC ATGTTTTGAA GCAGCAGGTC 540
 CAGGTCACTT TGTATATAGA ATTTGTCTGT ATTCAATAAA TCTGTTTGA GGNAAAAAAA 600
 45 AAAAAA AAAA AAMTSGAGGG CCGAAGCT 628

50

(2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1053 base pairs
 55 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

CTCTTTTCTG CATTTCAGAG GAGAGAGAG ATCTTGACAA ATGCACTCTG CTCTTGCCCT 6
 TGGCTGGGGA AGGTGGGAT GGAACCTCTG GGGCTGCTCA TCTTACTCTT TGTACAGAG 11
 5 CTCTTGAGAG GGCACAGAC GAGCTCTT GAGGGCTGG CAGGACCTC CCTGCAGGT 16
 TCTTGTCCCT ATGACTGCAAT GAAGCTCTG GAGAGGTGA AGGCTGCTG CCGCCAGCT 21
 GGAGAGAGG GCGGCTCTA GCTCTCTT AGCAGCTACA ACTTGTGCT GTTCTCTT 26
 10 CTGAGAGGT GAGTGGAG CACAGCATC ACAGAGATA CCGTGGCTG CACTCTCACC 31
 ACTAGCTGG GAACTCTCA ACCCATGAT GCGCTCTCT ACCATGTGA GAGCTCTCAT 36
 15 GGCAGTGAAG CTGATCTT GAGGAGCTC CTGCTGGAG TGTGCTGA CCGCTTGGAT 41
 CACCGAGATG CTGAGATCT CTGCTCTCT GGGAGTCTG AGAGTTTGA GATGCTCAT 46
 GTGGAGACA GCACTCTAG GAGCTCTT GAAGAGAAA TCGCTTCTC ACCACTTCT 51
 20 ATCTCTCTG TCTGCTCTG GATCTCTT ATGAGATC TACGAGTGA GCTCTCTG 56
 GCTGCAGCT GCACTGACA GAAGCAGG ACACATCAC CAGTGAAT GACTCTGGC 61
 25 CATGACCCAG GTATCAGT CCAACTCTG CAGGGCTGA GAGACCTG AAGGAGATG 66
 ATGGAGGAA AAGTCCAGA GAAGTCCAC CAGGACCA CCGAGCTGC ATACTTGCCA 71
 CTTGCTCACC AGGACTCTT GTCTGCTCT GCAAGAGAG TACTTCTCT GAACACTGT 76
 30 TCTCTGAGC CCGGAGAA GAGCTCTT GAGGAGTGG GAGCTGCTA AGAACCTC 81
 ACAACTCTG AATCTGAGC ATTTTAAACA CTACAAATA AATCCAGAC TGTCTATT 86
 35 NNNNNNNN NNNNNNNN AACGAGGG GGC 91

40 (2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1075 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 91

50 GCACAGAGCT GATCTCTCT TCTGCTCT TGAAGGAAA GAGGAGATCT TGAACAGG 6
 ACTCTCTCT TGGCTTGGT TGGGAGAG TGGCATGGAG CTTCTCTGGC TGTCTATCT 12
 ACTCTTTCT ACAGAGCTCT CAGGAGGCA CAACACCACA CTCTTCCAG GGTGCGGG 18
 55 CCACTCCCTC CAGTCTCTT GCGCTATGA CTCCATGAG CACTGCGGA GCGCAAGG 24
 CTGCTGCCG CAGCTGGAG AGAAGGGCC ATGCCACCT GTGCTCAGA GGCACAAT 30
 60 GTGGCTGCTG TCTTCTCTA GGAGTGGAA TGGAGTACA GGCATCAG AGATACCT 36

235

5 GGGTGGGACT CTCACCATTA GGTGGTGBAA TCTACAACCG CATGATGCGG GTTCTTACCA 420
 GTGCCABAGT CTTGATGGGA GTGABGCTGA CAGGCTCAGG AAGGTCTGCG TGGAGGTGGT 480
 GGCAGACCGG CTGATGZACC GGGATGCTGG AGATCTCTGG TTCCCCGGGG ASTCTGABAG 540
 CTTGAGGAAI GCGGAGCTGG AGGATAGCAT CTTCAGGAGT CTTTGGAAAG GAGAAATCC 600
 10 CTTCACAGCG ATTTCAGTCC TTTCTCTCTT GGGCTGCGATC TTTCTCATCA AGATTCTAGT 660
 AGCCAGGCGG CTCTGGGCTG CAGCCTGGCA TGGACAGAG CCAGGGACAG ATCCACCCAG 720
 TGAAGTGGAG TGGGGCGATG ACCCAGGGTA TCAGCTTCAA ATTCTGCCAG GCTTGAGAGA 780
 15 CACGTGAAGG AAGTGTATG GAGGAAAAGC CCAGGAGAG TTCCACCCAG GACCAGCCCA 840
 GCCTGCATAC TGGGCACTTG GCCACCAGGA CTCTTGTCTC TGGTCTGGCA AGAGACTACT 900
 20 CTGCTGAAAG ATTGCTCTCT CTGGACCTTG GAAGCAGGGA CTGCTTGAGG GAGTGGGGA 960
 CTGCTAAGAA CAGCTGACAA CTCTGAAATA TGGACATTT TAAACACTTA CAAATAAAD 1020
 25 CAAGACTCTC ATATTTAAAA AAAAAAAAAA AAAAAAAAAA CAGGGGGGGN CCGGG 1075

(2) INFORMATION FOR SEQ ID NO: 93:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2492 base pairs:

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

35

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 93:

40

TCCCGACTCA GCTTCCCAAC CTGGGCTTTC CGAGGTGCTK TGGCGGTGT CCCCACCACT 60

GCAGCCATGA TCTCTTTAAC GACACGCGAG AAAATTGAG TGGGATTAAC AGGATTTGGA 120

GTGTTTTTCC TGTCTTTTGG AAGGATCTTC TTTTGTGACA AAGCACTACT GGTATTTGG 180

45

AATGTTTAT TTTAGCCCG CTGGGCTTTT GTAATTGCTT TAGAAGAAG ATTCAGATT 240

TTCTTCCAAA AAGATAAAAT GAAAGCTACA GGTTTTTTTC TGGTGGGTAT ATTGTAGT 300

50

CTTATGCTT GGGCTTGGAT AGGATGATC TTGGAATTC ATGGATTTT TCTCTTGT 360

AGGGGCTTCT TTTCTGTCT TGTGCTTTT ATTAGAAGAG TCCAGTCTT TGGATCTCT 420

CTAATTATC CTGAAATTAG ATGATTGCTA GATAAAGTGG GAGAAAGCAA CAATATGCT 480

55

TAACACCAAG TGAATTGAA GATTCATTTA AATATTTTGG TATTTATAA AGTCATTTGA 540

AGAATATTC GAGCAAAACT AAATACATG AAATAGCTTG TAATGTTCTT TACAGGACT 600

60

TAAACGTAT AGCTTACAAA CTACCGGAG CAAATTAGCA AAGAGCAAT GAAAACAGG 660

	TTCTACTGTA GTGAACTAAG AAGAACTGAG GAACTAAACT GAGAGAGGTG AAATCCATGT	720
	TAAAGATGCT TAAGAACTG TTGAAGGCTA TTCTGTGTGT TTTTUCACAA TGTGGGAAAT	760
5	TCAGCCATCC TTAGAGAAGT GTGCTGCTGT TTCTTTTGT TTTTATTTCG AAGGTTGAG	800
	AGCATTCATA GGCATTTGCT TTTTAGAAAT GTCACTGGA ATGGCAAAAA CATTTTCAGT	840
	TGCACTGTA CTCTGGAAT GATGCATGA TTAATGGA TGTGTCAT TTAAGATAT	880
10	AAAACCAAG AAACCCCAAT TTGAGGCAAT GATTAAGTTT TTTTGTAAA CATGTTAAAT	920
	ATAAACTTC TGTGTTCTT CTGAATCTTA ATATTTCAA GGTAGGTGAA AATGTGAAT	960
15	AGATATCTT TGTGGAATA TGAAAGGTC ATTTTCTACT AACTTTTAGT TACCTAATTA	1000
	TAGCTAAGT TTGTCAGCAG CATACTCCG AATGTCTCAI ACTTCTTAGG AGTGTGCTT	1040
	CGTAAGTAT TGTCTATAT ATTCATTAG TGAAGTAT TAACAAAAA GATTTCTGT	1080
20	CGATGAAGA AGTATTTGT TCTAATCTT GGTTCATTGA ATAGTATTAT TGAAGATAT	1120
	AATGATGGA AACCAATGG ATTTTCTCA TGTATGATG TAATTTTTCT TGTCTTTT	1160
25	TTTTTTTAA ATTTTAGCAG TGGCTATTA TTTTCTTTC ATAACTTAA ATAACTTTA	1200
	ATAATCTTA CTTAAGACA TGTAACTAT TAAAGGTTA AACTTATGTC TGTTTTAAA	1240
	GGCTATTTCA TTTAATCTGA GTTTCTCTT ATTTTCAGCT TTTTCTTAGC AATAATAGT	1280
30	CATTAGCAI GACATATCTT TCATATGATC ACTATTTTG AGTTAATTAG AAAATACCT	1320
	AGTTACATG CTAAAGTCAT TCACTATAA TAACTGACT RIGTTTCTT AAGAACATGA	1360
35	CATTAAAAA AAGTGCTTT TTTTCCATG TTGCTGATTA TTAGACAGTA GGAATTAGC	1400
	GTTTCTTTA GTTTTACAAG ATGTGACAG TTTAGTCTA GATSTAGGGA AATATTGAA	1440
	CAGCCATAGT ACTATTGTT TTACCACTGA TGTACTGCT TTGTTTTTT AACATTTGCA	1480
40	AAGCTTTTAA ATGCAAAAA GTTAATTTA AATCTGAGT ACTTATTTAC AAACATCTC	1520
	ACAAAAATAG ACTACAGCTT ATTTTATTT TAGTTAAATC TCTTAATACA CAGAENAAT	1560
45	CCCAATCTTC CTATCTTAAA TAAGGAAABA CTTGTTGAT AGTGTGATGG TTTATCTTT	1600
	AGGATTAAGA CATTTTGTGT ACTTGCAATT GACTTAAGAT GTATCTGTGA AATTTGATG	1640
	ATATTGACAA ATGAGAGCTC CTACCTGAT AGTTAATGGA ATAATAAGAG GTTACTCTT	1680
50	TGCTAATAT TCTTCAAAAA AGTAATATCC TCACTTGAG AGTGTCAAT ACATACCTT	1720
	AGGATGACT CTATATAAGG TGCCCTTAG AATCTTTTA CACATATTTT TGACCAATAT	1760
55	TATTACAAAT GTCTTGATAA TCTACCTTT TTAGAGCAAG AATAGTATCT GTTAATSTAT	1800
	GGGACATCTG TATTTAACTC CTTGTAGAG ATGAATTTCT ATCAAAAIGT TCTTTGCACT	1840
60	GTACAGAGA TTCTTTTCTT CAATAATCTT AATTCAAAGC ATTATTAGGM CTGGAAGG	1880

TTTGTAAATC TCCCCSTCCT TGSTAAAGGT TG

2492

5

(2) INFORMATION FOR SEQ ID NO: 94.

(i) SEQUENCE CHARACTERISTICS

10

(A) LENGTH: 3058 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

15

ACCCCTAAATC AACAGACAAT GGCATTGTCTG AAGAGCAACC TGTTAATGAA ATCATGTTAAT 60

AAATCAAGGT TTGGCTTCAG TTTAAATCAC TTGAGGTATG AAGTTTATCC TGTTTTCCAG 120

20

AGATAAATAA AAGTTGATCT TCCCAAAATA CCATCATTAG GACCTATCAC ACAATATCAAT 180

TAGTTTCTTT TGTTCSTTTS TTTTCTGTTT TTTTCTTGG TAAAGCCATG CACCACAGAT 240

25

TTCTGGGAG AGTTGAGAGA CAATGGTCTCT GACATAATAA GGATCTTTGA TTAACCCCTC 300

TAAGGCATGT GTGTGTATAC AAATATACTT CTCTTGGCT TTTCGACATA GAACCTCAGT 360

TGTTAACCAA GGGAAATAC ATCAGATCTG CAACACAGAA ATGCTCTGCC TGAAATTTCT 420

30

AAGATGCTTA GSACTCACCC CATTTATCCA GGTCTTTCTG GATCTGTCTA ATCAATAAGT 480

CCATATAACA CTGTCTAAAC ACTGGGCTCT ATCAGCCAGG GATAAAACA GAGATCAT 540

35

TCTTGGAGCT CCTGCATCAG CCIATTCAAA ATATCTCTC TCTCTAGCTT TCCACAAAT 600

CTAAATTTCC TGTCCCAAGC CACCCAAATT CTCAGATCTT TTCTGGAACA AGGCAGAATA 660

TAAATTAAT ATACATTTAG TGGCTTGGGC TATGCTCTCC AAAGATCCTT CAAAAATACA 720

40

TCAGCCAGC TTCAATCACT CACTTTACTT AGAAGAGAGA TATAAGGCCC TGGGATGCA 780

TTATTTATC AATACCAATT TTTGTGGCCA TGGGAGACAT TGCTAATCAA TCACAGCACC 840

45

ATTTCCATTT AAGCCCACTG ATTTCTTCACT AATCTTTCTC AAATTACAAT TCCAAAGAG 900

CGGCACTCAA CAGTCAGATG AACCCAACAG TCAGATGAGA GAAATGAACC CTACTTGCTA 960

TCTCTATCTT AGAAGGCAAA AACAAACAGG AATTTCAGG GAGAATGGGA AAGCCAGGG 1020

50

GCATAAAAGG TACATTCAGG GGAAATAGA TTTAGGCAGA CTGCCTTAGT CAGGSAACAT 1080

GGGCGCTCAA TCTGCAGTGC CAACACCAAA CTGACACATC TCCAGGTGTA CCTCCAAACC 1140

55

TAGCCTTCTC CCACAGCTGC CTACAACAGA GTCTCCAGC CTTCTCAGAG AGCTAAAACC 1200

AGAAATTTCC AGACTCATGA AAGCAAGCCC CTAGCTCTC CCCAACCCTG CCGGATTGTT 1260

TAATTTTAA AACTACTAGG TCTTCTCTTC ATATATTTCC TCATAAGCAG GAGCCAGAA 1320

60

ATTTCACTCA CTTGCAGTGA CATTGCTGSA CTCTGAAAA CATTCCATA GGAGAATGCT 1380

	TTCTGAGGCT TCAAGTGTGA GAGACACTGA GATGATCAGA ATTGTTTGA GTGTGGGAG	1440
5	TCTAAACAG TCCACTCAGT GGTGCGGCT GGGGGTGA TGGGGTGA TTGAAATGT	1500
	GATGAGGA TGCTGAGCT GTGAGGAA AGATGGGAG CATGATGGA AATGCAACAT	1560
	TGATACAGT ATTGATATG GATGCGGCT TCAAGATGT TATTGAGGA AAGCATTGT	1620
10	ATGAAAGAG GAATGAGAA TGTAGAGA TTTGAAAAG GTTCAGGAG ATTGTTTGT	1680
	AAATGACTG ATGTGTGAT GGGGGTAT TTGAATTA AGTTAAGAG GAGAAACTT	1740
15	TATAGAGTG TATGAGAG GTGATGCTT CATAGTGGG TATTACAGG AGGAAATGT	1800
	TTTATGCTT TTACAAAGT GATGAGAGT GTGTGCTT CTTTAAAAG GAGGAGAGA	1860
	TGTGATATG ATCAGGAAAG TGTGAAAG TATTGTTTG AGCCCGGCTG TATGTGCTT	1920
20	TTGAGAGCT TTTTGTGTA TTAAGGAA ATGTGTGTG TATATATGG TATTCATGT	1980
	GTTAGAGGA AGATGCTT ATCAGTGTG AAAAAAAG ATAGTGTAG TAAATATTA	2040
25	TAAAGGCGT GATGTTGAT GATGAGTAT TGTACCAAG GTGTGCTGT GTTTTTCCT	2100
	ATGATGAT TGTGTTTATA ATTATGAA TAGTACTGA AATGAGAGA CCGTGTGTT	2160
	CACAGATTA ATAGAGGCT TGTAGAAAG CATATCTGT TGACAGGAG CTTACAGTT	2220
30	CTTGTGAG GTTGTGTGA CCGTGTGTG AGACACAAG TGTGTGTT ACCAAATGT	2280
	AGAAGAGAG TGTGAGTA ATAAATAGT TTGATATCT GCGCATGGT AACCTGATG	2340
35	TAAATATAT CAGATGAGT AGATGATG TTGAAGTGT ACATACATA AAGTCAAG	2400
	ACTATGAG ATGCGGTAA ATGATGAA AGAAGAGGA AAAATATTA TAGATGATA	2460
	AGCAAGGCT TCAGGCAAT GATAGGAG TAAAAAAG AATTAATCT CCGTAAATG	2520
40	GTGAAATAG TTGTAAGCT ATATAAGAG ATATGCAGT AAAAGTCT TAATGCACAT	2580
	CGTGTGAG TGGAGTGT TAAGAAATG CTTTCTGT TTATGAGG TCTGATAT	2640
45	TATGATAT AGATAGCA ATAAAGAA TTAGAATAT ATTGTTATA CACTTAACAT	2700
	TAAATCTCT TAATGCTT CTTTGTGA TAATCAGAA GATAGTATG GATGTTCAAT	2760
	GCTGTGAT GATGTTATA AAAAGGAG TATCACTATA CCGTGTAT GAGACTGG	2820
50	CAAAAGGCT AATAGAGAA CCGTGAAGT TGCTTTTTT AAAAAAAG TAAATGCT	2880
	AAATCAACT TTTTGTGG TTGCTGTT GTTATAAGT GCAAGGAT CAGTCTCA	2940
55	ATATCTGAT CATATAGCA TGTATAGGA GATGGGCA AACCTGTCA ATGACAGCT	3000
	TGGAAGTGT TTTTAAAA AAATATTA TTTTAATCC AAAAAAAG AAAAATG	3060
60		

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1099 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95:

10 GGGTTTGTTAG CTGCTTCGGCA GGGGAGGCGG GGGGAGGCTCG CAGAGTCTTA GGGGTTGGG 60
 GGCNTCTCTG CTCTCCCTT CTGGGCGCTC GGGGCGCGG CCTCCGCGGT GCGTGCCTTC 120
 15 GGTCTCAGGT TGAGGAGCTT AAGTTTGGGA AATGTTGTG CATTCCTTGT ATGTCATTTC 180
 CAGTTCTGCT CTGATCTAC AAAAAATTCC TGGAGCCATA TATATACCTT CTGGTTTCCC 240
 CTTTCGTTAG TGTATATGT CCTAAGAAAG CAATACAAGA ATCCAAATGAT ACAAACAAG 300
 20 GCAAAGTAAA CTTTAAGGGT CCAACACAGA ATGATTACC AACAAAAGGA CCAACAGAA 360
 TCTGTGATAA AAGGAAAGAG TAAAGAAATT TTCTAAAGG ACCCATCAT TTAATAAATG 420
 25 GACCTGATAA TATGAAGCAT CTTCCTTGTG ATTTCTCTTG ACCTTTTAT CTGAGACCGG 480
 AATTCAGGAT AGGAGTCTAG ATATTTACCT GATACTAATC AGGAAATATA TGATATCCGT 540
 ATTIAAAATG TAGTTAGTTA TATTTAATGA CCTCATTCCT AAGTTCTTTT TTGTTAATG 600
 30 TAGCTTTCAT TTCTGTTATT GTTCTTTGAA TAATATGATT AAATAGAAGG TTTGTGCCAG 660
 TAGACATTAT CTACTAAAT CAGCACTTTA AATCTTTGG TTCTCTAATT CATATGAAT 720
 35 TGCTGTTTGC TCTAATTTCT TTGGGCTCTT CTAATTTGAG TGGAGTACAA TTTTGTGTG 780
 AAACAGTCCA GTGAACTGT GAGGGAAT GAAGGTAGAA TTTTGGGAG TAATAATGAT 840
 GTGAAACATA AAGATTAAAT AATTACTGTC CAACACAGTG GAGCAGCTTG TCCACAAATA 900
 40 TAGTAATTAC TATTTATTGC TCTAAGGAAG ATTAAAAAAA GATAGGGAAG AGGGGGAAC 960
 TTCTTTGAAA AATGAAACAT CTCTTACATT AATGCTAAT TATAAAATTT TAATCCTTAC 1020
 45 TGCATTTCTT CTGTTCTTAC AATTTATTA AACATTCAGT TTAATCGGTA AAAAAAAAAA 1080
 AAAAAAACC GGGGGGGG 1099

50

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1580 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:

60

240

GAGAGAGAT AGAGCTCTT TTCTGAAA ATTCCHYCT TTAATTCG GTTCGAGAA 6
 GTTCAGTTTC TTCTCTGAC CAGCAGCTT GATTCCTTT ACCTTCCTT GTTCGAGAG 12
 5 GTTCCTCTTT GTTCCTTCA CAGCAGTAT GTTCGAGAG CAGCAGCTT CAGCCTTAT 18
 AGTGTGAGC AGATAGAG AATTAAAA GAGAGAGAG AGAGCTTA AAAAAAGAA 24
 10 AATGAGAG AGGAGAG TTCTTGGCA TCTCTCTT TTAGCTAG GAGAGCTT 30
 TGTCTCTCA GAGAGAG AGCAGAGCTT GAGAGAG GTCTCTGAA GAGAGCTTA 36
 CGATCTGCTT ACTGAGAG AGTCTCTAC CAGCAGCTT CAGCCTTCTT GTTCCTCTT 42
 15 AATCTCTTAA TCTGAGAG AGGAGAG GTTCCTTAA TCTCTCTT GTTCCTTAA 48
 AATGCTGAG CAGCAGAG AAAAAAAA GTTCCTTAA TCTCTCTT AAAAAAGAT 54
 20 CAGCAGCTT GTTCCTTAA CAGCAGCTT TCTCTCTT TCTGAGAG AATCTCTT 60
 CAGCCTTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 66
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 72
 25 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 78
 CAGCCTTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 84
 GTCTCTCTT CAGCAGCTT AAAAAAGAT TCTCTCTT TCTCTCTT GTTCCTTCT 90
 30 CAGCAGCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 96
 CAGCCTTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 102
 GTAGAGCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 108
 35 GTAGAGCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 114
 ATCTCTCTT ATCTCTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 120
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 126
 40 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 132
 ATCTCTCTT ATCTCTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 138
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 144
 45 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 150
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 156
 50 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 162
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 168
 55 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 174
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 180
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 186
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 192
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 198
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 204
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 210
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 216
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 222
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 228
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 234
 TCTCTCTCTT GTTCCTTCTT TCTCTCTT AGATCTTAA TAGAGCTT GTTCCTTCT 240

(2) INFORMATION FOR SEQ ID NO: 97:

60

241

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 678 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 95

10 ATATTYTTTT AGTTAATNT CCAAGATACA GGTMTGAGGA GCGAGCTATG TTTAATGAGG 6
 GCTCTCTTGT TTCTTAGAGA TGAGAGAAAT GTATACTAAT CATTNTAATT TGTACTTAAA 120
 ATACATTTTA CTAATCATAT TGATTTTAAA TATGACAAAT TCTTCTAGTA GATACTAATC 180
 15 TTCTTGTMT ATCATATTGT CCAAGAGAGG CCTAGGTAAA AATGGGTTCG ACCTAGTCTG 240
 TTTGTATAAC ACCTTCCGCG GTTCCCTCTC CATCCCTGCC AATTGGGCTG TATGCATATT 300
 GACAAGCAAA TAAGAAAACC TTAGSTTTCT TGTATTGAA TTCCAAAAC AATAAAAGGT 360
 20 TTGACTCAA GATTTGCATT CAGAAGAGG CAGAAATTTT GTCTATCTT TTTATCATT 420
 TGTGAAGTGT TGTCTCTCTG TATGCTTAGA AAATTTTACA CACAAGGAAT GTTTGAAAAA 480
 25 GTGAGAATTT TAGAGTGCTT GGTGCTTTT TATTTGGTCA GTGCTGAGCT GTIARGTGTI 540
 TAGGGAAATA ATGCTTCAGG ACCTTTTGA CAACACAGYT TCATGAATGA CYGGGGGATA 600
 TTWAGTGTGT GCTGACAAAA GCGAGGGAGT GGGCAGTGG AATGGGGGAC CCTTACCATT 660
 30 GGAAAACATG CATTCTGT 678

35

(2) INFORMATION FOR SEQ ID NO: 98:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1253 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98

45

ACCTCCCTCC CTCTCAGACT GGTCCGAATC CACGCCTAGC CCAGTCACTG CCACTGGGGC 60
 CATGGCCACC ACCACTGAGG CACTAGCTTG CCAGTCACTT CCCTTGTCTG TTCCAGCTC 120
 50 CCTTGTCTAG GCTCAGATCC AGCTGGGGCC CCACCGGNA GTTACCCCA AGAGGCAAGT 180
 NTTGGCTTGA GAGGCTCTC ASTTCTTAGA TCTTGGGGCC CTAAGAGAC CCCCCTCTG 240
 CCTCCTTTCT TTCTCTCTCT CTTCCTTCTT TTTATCTTTT TCCATCTCTT TCTCTTTCCA 300
 55 CCAACCCCTC TATATCTTCT CTTTCTCTCT TTTATCTTTT TCCATCTCTT TCTCTTTCCA 360
 CAGTCTTCTT TTATTATATA TATTTGGGGG CTACCACTCA CCCCCTCTCA GTCTTGTGAA 420
 60 GAGTCTGAGA CCTCCTTCTT CTTCACTCTT CTCTTCTCTT ATTCTTTCTT CTCTCCTTCT 480

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55

GGTCTCTGAG TTTCTTACAG TCTGACATGA AGGAAATATG ATATATTTTC TTTTCTTT
 TTTTCTTACA TTTTCTTACG AAGCAATCTG AGTTAAAGCA AGGCAATATG ATATATTTT
 ACAAATATTA TATATAGAGA TGTCTCTTCC CTTCTGTGAG CCCCCASTGT CCCCCTGGG
 TGNAGTCTGT GATCTGATTC GGGCAAGCTG GATCTCTGT AGCTAGTACA CAGGCATGA
 TGGGATCCCG TGTATCGAGT ACAGGATCCA GGTATCTACC AAGTAGGCAC CTTTGGGG
 ACTCACTGAG GTCAGAGCTG GGGGAGTGT TGGGAGGCTG CTCTCCACCG CACTTCCCT
 ACTTCACTAG ATTCTAATG GGAATCTCTG CTAAGCTCTG CTTGGGAAGG GCCCCTGTG
 AACTCCCTGT GGCAGAGCTG CACTCTTGGT CATCTCCCTG TGGGAAGTAS GGGGCTGTG
 GTGGGAATG GGAATCTCTG CTAAGCTCTG CTTCTCTTCA TGTCTCTGTA AATCTGGGA
 TACAGAGAGA GGAATCTCTG CTAAGCTCTG TGTCTCTTCA GGTATCTCTG TGGGAGAG
 TTAGGAGAGA ATAGAGTAT TTTAGAGCTA TTTTCTTAT AAGGCTCTG GTCAAAATG
 CTGTCTAGT GATCTCTCA GGTCTCTAT GTACAGGCTG CACTCTCTCT CACTCTCTA
 TAAAGGAATA GTTACACTG AAAAAAAAAA AAAAAAAAAA ACTTGAGGGG GGG

540
600
660
720
780
840
900
960
1020
1080
1140
1200
1260

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 447 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 99:

40
 45
 50
 55

CAAAGATGA AATTACCACT TCTCTCTCTC TTGGCAGCTG TAGCAGGGGC CCIGGTCTAT
 GCTGAAGATG CTTCTCTGTA CTGACGGGT GCTGATCTG CCGAGGAAGC TGGGAGCTG
 AAGGCTAATG AAGGATCTG AGTCCAGCA GAACAGTTT CACCCGAGA GACAACCA
 AGAGCCGAGG AGATCTCTG GGTAGCAATT CAGGAGAGG CCAAGGTCAT CTCAAGAG
 CAGGACATA ACCCTCTGAA ATCATATG GAGAAAGTA TCTTACTAAC AGAACAGG
 CTGCAALAG CAGGAAAGG AATTCACGGA GGTCTCTGAS GTGGAAACA ATCATCTGA
 AATGAGAGT AATTCACGA AATTTACTG AAGAAATICA CTCTATTAAA ACCATGGGA
 TSAGAAGCTG AAAAAATKG GATCAT

60
66
126
186
246
306
366
426
447

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 611 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:

10 GGTCTGGGSA GGTGACATGT TGGGCTGTGG GATCCCAGCG CTGGGCCTGC TCCTGCTGCT 60
GCAGGSWTNG GCAGACGGAA ATGGAATCCA GGGATTCTTC TACCCATGGA GTGTGTAGGG 120
TGACATATGG GACCGGAGAG GTGTGTGGGG CCAGGCGGCC ATTCGATAGC CCCAACYTC 180
15 GCCTGCGTCT CCGGTGCTGC TACCGCAATG GGTCTGCTAC CACCAGCGTC CAGACGAAA 240
CGTGCGGAGG AAGCACATGT GGGCGCTGCT CTGACCTGC AGCGGCCTCC TCCTCCTGAG 300
20 CTGCAGCATC TGGTTTCTTT GGTGGGCCAA GGTGCGGGAC GTGCTGCATA TGCCCGGTT 360
CCTGGCGGGT CCGTGTGACA TGTCCAAGTC CGTCTCGCTG CTCTCCAAGC ACCGAGGGAG 420
CAAGAAGACG CCGTCCACGG GCAGCGTGCC AGTCGCCCTG TCCAAAGAGT CCAGGGATGT 480
25 GGAGGGAGGC ACCGAGGCGG AAGGGACGSA GAGGGGTGAG GAGACAGAGG GCGAGGAAG 540
GGAGGATTAG GGGAGTCCCC GGGGGACTGG TCAATACAGA TACGGTGGAC GGAAAAAAA 600
30 AAAAAAAAAA A 611

35 (2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 609 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 101:

45 GCATTGCTAA AGCTGGCACT TGAAACCAGT TGGACGGCCC AGCTTGCGTC TCTTCTGCC 60
GAGTGGGCCT CTCAGTACAC TCGTGCCCTG CTGGAGGACA GAGGGGCACC TCAGCCGCC 120
CCAAGTCCAG AGCAGAGCAA TAAGGTGCGC CTGAGGAGC CCGGTGCGG GTGGGGGTG 180
50 GGGGCGCAGG ACCCTRARAT GCCACCAGSA CCGATGCGC CAGGAAGGGC GTGGACATG 240
AGGCTGTTTT TACAGTTTTT TTTTTTTTGT TGTCTGTTT TTAAAGAATA CAGAAGGAG 300
55 CAAGCTTTTT TGCACTTTGT ATCCAGCTGC AAGCTCAGGG CAGAGTCAAG GGCTTGGGT 360
GGAAAAAAGT GACTCAGAGG AATGCATAAT TGACCCTTGC AGCTACCCAA TAGCCCTTG 420
AGCTGGCACT GAACCAGGCT GCAAGATTG AGTGCCTTAA AAACACAAGG CCTCTAG 480
60

CTG-CAGGGA TGTCCCTGTG CCGAGCATTG GAGGCTCGAA GACTGGTTTC TAGCATTACG 540
 GGTACAGGTC ATGTGTCTCT AGAAGGCTCC AGAGATTAT TTACCTTGA GTGCATTTT 600
 5 AATGTTCT 609

10 (2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1770 base pairs
 (E) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 102:

20 ACCGTCGGA ATCCCGGCTC GACCCAGGCG TCCGGGAAAT TGAAACTGAG TGGCTCAGG 60
 TGGGAGAGAG GAAAAGGCGA GGTATACAGG AGGCTCTGCG CTGAAGGCAG AGGTAAACAT 120
 GGGCTTCGGA GCGACCTTGG CCTTGGGCTT GACCATCTTT CTGCTGTCTG TGTCTACTAT 180
 25 CATCATCTGG TTACCTGCTT CCTGCTGCTG CCTTACAG AGTGGCCGCC GAGCAAGTCC 240
 GGTGTCTGAC ACCACACAT CCACCACTGT GTGCTATGCC CTTATCTCT AGCTTCAAG 300
 30 TGTGCGGCGC AGTTACCTTG GACCAAGCTA CCAGGGCTAC CACACCATGC CCGCTCAGCT 360
 AGGGATGCGA GAGGACCCCT ACCCAATGCA GTACCCACCA CCTTACCCAG CCGAGGCCAT 420
 GGGTCCAGCG GTTATCCAGG AGACCTGCGC TGAGGAGCA GCGCGGCTT ACCCGGCCAG 480
 35 CCAGCTCTCT TACAACCGCG GTTACATGGA TGCGCGGAGG GCGCTCTCTG AGCATTCCTT 540
 GCGCTCTCTG GTTGCACCTT GTTATCTCTG TGTGTCTGCG TGATGCTGT GCGGCTCTG 600
 40 TTCTTACCTT CCGATGCTCT GTGTGTCTCT CTTGCTCTTA TATGTGCTT CTTTATAT 660
 TGACAGCTTG GGAACAATC CTGCGCAGAG TGAGCTGCGA CCAGACTTCT TCTCTTCTT 720
 CACTTGAAT TATGCTTCTT AAAATCTTAA GCGAACTCA AGAATGCGG TGTGCGCGG 780
 45 CACCTCTTGA GTTGGCGCTT GAGAGCTGCG GGTCTCTTCA GGCACATCT GGAATTTTCT 840
 TCGATTTTAC CTTAGCTTGA CCAATAGCG CCGGCGACAG CAGGCTGCGG CAGTTTCTG 900
 50 TGTATGCGAG ATGTGCTCTG GTTGTGAGAG CCAAGCAGAG TGTGCTTGA GCGATGCTT 960
 CCGTCCCTGA CTGGGGGCTA CCGCTTCTAG AGCGAGGAGC ATGATGCGAG CCAAGTTTCT 1020
 GAGCTGCGCA AGTGGAGCTT TATCTCTTCT GCGATATCTC CCGATGCTCT CTGGAGCTT 1080
 55 TGTGCTCTGT TGAGGATGAG GCGCTCTCTT GATGCTCAGG CACCTCAGGC AGAGCCCTAC 1140
 TCAGCTCTAC CTGTGTGCTT GAGCTGCTCT CTGTGCTCTG ATTTGCTCT GAGACCTCTG 1200
 60 GAGGCTCACA TGACACACA GGTATGCTGC CCGAGGAGAG CTGTGCTCTC CTGTGCTGCT 1260

CTGCGCTTCC CACAGGTGAG CAGGGGTCTT GTTCACCAAG AACTGABTT CTCTTCCCTG 1320
 CAGTGTCTTC ATTTTATTTT AGCTAAACAT TTNGCTGT TTCTGTTC AACTATATAG 1380
 5 TTGATATGAG ACTGAAACCC CTGGGTCTTG GAGGAAATTT GGTTCAGAGA TGGACAACCT 1440
 GGCAGTCTTG AGTCTCTGCT TTGAGACAT AGCTGATGG AATATGCAAC AACTCTCTGA 1500
 10 CCCCAGTCCA CGTGTCTCTG GAGGAGGGA CAGCTGAGCC AATGGGCGAT CTGGACCAAA 1560
 GGTGGGCTGT GGGGCTCTGG ATGGAGGCTT TGGCCAGAG ATGAATACCT CGTCTTCTT 1620
 CTCCCTCTAT TACTGTCTCA CCAGAGCTGT CTTAGCTCAA ATCTGTCTGT TTCTGTAGTG 1680
 15 TAGGCTCTGT AACTGTCTTT ATAAATATG CAATCGTTTG GAAAAAAAAA AAAAAAAAAA 1740
 TGTAGGGGGG GGGCCGTACC CAATGGCTA 1770

20

(2) INFORMATION FOR SEQ ID NO: 103:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1832 base pairs

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 103:

TGTGCTGAC GTCATCTGGA GAGATTTGT TTCTTTTTC TCCAAAAGGG GAGGAAATG 60
 35 AACTGCACT GGGCCACGAT GGAAGAGGG GAAAGCCCAG GGTACAGGA GGTCTCTGG 120
 TGAAGGAGA GGTAAACATG GGTCTGAGG CAGCTTGGC CGTGGCTGA CCATCTTGT 180
 GGTCTCTGT GTCATATCA TCATCTGCT CAGCTGCTCC TGCTGTGCT TTACAAAGAG 240
 40 GTGCGCTGA CCAGCTCCG TGTCAACAC CAGCACATCC ACCACTGTGG TGCATGCCCC 300
 TATCTCTAG CCTCCAAAGT TGGGCGCAG CTACCTGGA CCAAGCTACC AGGGCTACCA 360
 45 CACCATGCTG CTTAGCCAG GATGCGCAG AGCACCTAC CCAATGCAT ACCACCACT 420
 TTACCCAGG CAGCCCATGG GTCACCGGT CTACCAAGAG ACCCTGGCTG GAGGAGCAG 480
 CGGCGCTAM CCGGCGAGG AGCTCTCTA CAACCGGCT TACATGGAT CCGGAAGCG 540
 50 CCGCTGAGG ATTCCTGCG CTCTTGTGT GCACTTGGT TATCTGTGT GTCTGCTGA 600
 GTGTCTTCA GCGCGGCTT CTACGCTG ATGTGTGCT TGTGTCTCA GGCAGGCTT 660
 55 CTACGCTG ATGTGTGCT TGTGTCTT GCTGTATAT GTGCTCTCT CTGATGCTGA 720
 CAGTGGGA ACAATCTTG CCAGAGTGG CTGGGACCAG ACTTTGTCT CTCTCTCAG 780
 TGAAATATG CTCTTAAAA TGTCAAGCA AATCAAGA ATGGGTGCT GGGGCGCAG 840
 60

CTCTGAGGTTG GGTCTTGAGA GTTCTGGGCT TCTCTAGGCT ACATTTGAG TTCTTCTCCA 900
 GCTTACCCCTA GGTGAGCCA GTTAGGNTS TACACCCAGT GTGGTGGAST TTCTGTGTGA 960
 5 TGCAGATGTS TCTTGSTTTC GGAATGTTAS CAGGTGCTG CTGAGAGCCA TGGCTGCTCC 1020
 CCGGAGTTGG GGTACGCTT TGTAAATGTA GGGACATGAT GCAGAGGAGG YTTGGGATCT 1080
 GGTCAAGTTS GACTTGTAT CTTCAGGTA ASSTCCCATI GCTCTCTGTA GCTCTCTCAT 1140
 10 CCTGTTGGGG ATCAGGTCAG CTCTGATGC CAGAACACCT CAGGACAGAG CTTACTCAGT 1200
 TGTACCTGTT TGGCTGAGT GTCCTCTTTC CCGCATCTC CCTTGGGACC AGCTGAGGGA 1260
 15 CCACATGCAC ACAGAGTCTA GCTGTCCTTA GGGAGCTCTG CTCTCTCTTC TGGGCTTGGT 1320
 CTTCCTACAG GTGAGTAGGG CTCTGTCTCA CAGGCACACT CAGTCTCTCT CCTTGCAGTG 1380
 TTTTCAATTT ATTTAGGCA AATATTTTC CTCTTTCTG TCTTAAACAT GATAGTTGAT 1440
 20 ATGAGACTGA AACCTTGGG TTTTGGATGG AATTTGGCTG AGAGATGAG AACCTGGCAA 1500
 CTGTGAGTCC CTGTTTCCCG ACACAGGCTT CATGGAAAT GCAACAACCT CTGTACCCCA 1560
 25 GTTCACGGTG TTCTGGGAGC AGGACACCTI GGGCCAATGG GGCATCTGGA CCAAAGGTGS 1620
 GGTGTGGGGC CTTGATGGC AGCTCTGCC CAGACATGAA TACCTCTGT TCTCTCTCTT 1680
 TCTATTACTG TTTCACAGA GTTGTCTTAC CTCAATCTG TTGCTTTCTT GAGTCTAGGS 1740
 30 TCTGTACACT TGTTTATAAT AATGCAATC GTTTTGGAAA AAAAAANAA AAAAAAAGG 1800
 GSGGGCGCTC TAAAAGGATN CCCNAAAGG GG 1860

35

(2) INFORMATION FOR SEQ ID NO: 104:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2237 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

45

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 104

AGTTCCTGGT ACCTTATTTAC CAGCTCTGGC ATCTTACCA GTTATGAGAT TACTCACTAC 0
 50 CAGAAATGAG AAAATTGCTT TGAATATGTC TGGGAGTGC ATGATCTCTT ACATTACAGT 120
 TACTGTAAAG GATCTGAATG GCATAGCTT AACTCTCTG CAGGATACTC CTGTGAGCTT 180
 AAGAAAAGAA GATACATATG TTCAATTTTAA TGTGACATTT GAGCTTCAGA AGCATGTTGA 240
 55 AAAATTAACC AAAGGTGAG ATACTCTCTT TGAATTCAAA CACTACAAAC CTAAAAAAG 300
 GTTTACCAGC ACCAAGTCTT TGTCTTCTAT GGAGATGAGT GAAATTAAC CTGGGCTCAAT 360
 60 TGTAAATAGAA CTATACAAGA AACCTCACTG CTTTAAAGA AAGAAATGTC AATTACTGAC 420

	CAAGAAACCA CTCTACCTTC ATCTACATCA AATTTTGCAC AAGGAATGAT CCTGACATGA	480
5	TGAACCTGGA ACTTCTCTGA ATTTTACGAC TGAGTAGAAA CCATCATAGC TCTGTGTAG	540
	ATATTACCCC TTCAACAGGC AGGAAGDAAS CCGTACCCAG ATCAGTAGGC CGGACGGAGC	600
	CAATNGCHAA GTCTGACGAC AGAATTEAGA GTCCAGGACA TCACACTGAC GTATAGGAGC	660
10	CCTTGGGATA CAGGTCTTAT CTAGATTTTC AAACATGTTT TTACTTTTCT ATTAATTGT	720
	CAATTAATAG TCTATTTTCT AATTTACCACT TACCTCTACC CTGCTTCCTG GAACAATAC	780
15	GTGTGCGTA GATCTGCTC ATCTTCAGAC TTAATACAGC AATAAGAATG TGCTAGAGT	840
	TACACATCTG TTCACTTTTC CTCCAAATAG CTCTTTTGAC TTAACGTCAA GCTTTGGGT	900
	GATGTGCGTA GGGTACTCTC AAATGCTTTT GAGAGGAATG GACCCAGTTC TGCTGCTAA	960
20	GAAGGTCTCT CTGAGCTTTT ATAGGAGGCA CCTCTGAAGT GGGCTAAATT CACCTTGAT	1020
	TGATAGTTTT CCGCTTAGA AAGTGTGCTT TGCCAGATC AGTATCCAC ATGGGAGTCT	1080
25	TCCCTAGGTT CTAGTCTGA TTGTTTCCAG ATGACCAGAT TGTCTTCTG AAAATGAGCA	1140
	TATTTTTAGT CATCTGATG AGCTGTTCTT CTACATCACA TTGTTACTCT TTCTGATGAT	1200
	GATCTAGGG TTAACATGCG AACCATCTCA AATAATTAC AAAGTTTATG ATGGGTTTAT	1260
30	AAATGCTTCT AAACAATCTA ATCTAAAAAT AATGAGTCA GATGCTAACG AGATACTGCA	1320
	GGCATAACTG CTGTTTCTCT GACAACGAT TGTGAAACCT TAAACCTGC ATACCTCTT	1380
35	TTACAGTGAG GAGTATGCAA AATCTGGAAG GATATTCTAT TTTTCTTATA TAGGTAGATA	1440
	GGATCGCCAT TTATTTCTA TTIAGATATA CTGACATTC TCCATATGAA AATATGCAG	1500
	TCATTAGCTT ACTATAATT ACTTTTGACT TAATGCGGCA TAAATAAAAC TTTCATAGTA	1560
40	CACATGAGGT GATATTTGA TACACAGAAC ATTTGCGGTG GCTTTCTGT GGGTTAGAT	1620
	TAAAGCCAC ATATTTTAAT ATTCATATT TTAATGAGC AATGCATGAG GGAATGCAG	1680
45	TGTCAGTACC TGCCCTATTI TAAACTAGT GTAATCACC TAGTCATACC ATTCAGTAT	1740
	TTTGCTTTTT AAATAAGTA ACCACAATTA AGTTGTTGTA GGCCTTGAC TTCAAGAGAT	1800
	CTAGTCTTTA CTTCAGTAT TCTCTTAGGT CCATCTGTCT TACTAGACGG ATGTTAATAA	1860
50	AAACTATGCG AGCTGGAATG AATTCTCAGC CAAATTTAGT CTCTCTCTC ATCTTGATT	1920
	GATTAATTCC AAATCTAAA ATGATTCACT CCACAATAGC TCTAGGGGAT GAAGAATTTC	1980
55	CCTTACTTTG CCGCTTCCT AAGACTGTGA GTTGTCAAAT CCTAGACTG TAAGCTCTTC	2040
	AAGGAGCAAG AGGCGCATTT TCTCGCTCTC ATGTAATTTT TCTAGGTGT TTGGCAGCAT	2100
	TCTGTACCTT GTGAGTACT CAGTACCTTT TTTTGTATGT TCTGACAAAG ACCTGAAAA	2160
60	AAATCCCTTA AAAAAAAAAA CCATTAAAGT GTAGCAAAAC CCAAAAAAAAAA AAAANAAAAA	2220

ACTCGAGACG GGCCCGC

2235

5

(2) INFORMATION FOR SEQ ID NO 305:

(2) SEQUENCE CHARACTERISTICS.

(A) LENGTH: 1822 base pairs

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:

GSTCGACCCG CGGTCCGGA ATTTCGTAG CAATAAGTTT GTGCATGTAT AGTAATTTC 60

ATTAAGCAAGG TTCTAAGCTC TGCCCTCTGG GTTCAAGTGA TTCTCGTGCC CCAGCCTCC 120

GAGTAGCTGG GAGTACAGGC ACGTGCCAGC ACGCCAGCT AATTTTATA TTTTAACTAG 180

AGACGGGGTG TTGCTGTETT GGCCAGGCTG GTCTCAAACCT CCTGACCTCA AGTAATCCAC 240

CTGGGCTGGT CTTTCATGT CTTAACATGG CATGTCCTTT AGTTTCATTA TTTTCTACT 300

CCTTCTATGT CAGGAAATTA CATTTTGCAT GTCTTATGGA GATGCTGTTA ATTGCTTCAG 360

TGAGTGCTT TCGAATCTGC AGACCATTTA CATTTCTGT TTGCAGCATG CTGTGTGCAA 420

ACACTCAGTA ATTGGGAGTA TTCAATTATG TGTIAGGGGT CTTCCTATTT CCAAATGTG 480

TGAATTGTCT ATTGATGGGA TTTTCAGATC TTTTCATGAG AACTGGAAAT GTAGCTGGGT 540

GGCAGCTACC TAGGTTGCTA CGTAGTGAGT AGACTTTCTC TTGGGTATAG TAAGCCTCA- 600

ACAGCTTCA CTTTATCTA CTTTACTGT GAAATAAAA CAGTCATTTT GTTCTGAAA 660

ATAAAGATAG CTTTCTGTAG AGAAGGAATT CCTACCTCTA AAAGCTGCCT TGAGAACTCA 720

GAAGTGGCAG TTTCTGAGG TGATTITAA ATTTCAGTAT TAGGSAGAGT CCAGCATTTC 780

CTGACACAGG TTATACATAA CTAATGTATG ATAGCAAATG CAAACTATT ATAATGTGG. 840

GTATCTTGCG CATACACAGG TTAGAACAAG TAGACTCTGG CAGCAGATCT CCAGAGACCT 900

AAGTTTAGGT TCICATAGTG TATTGAAGT AATTATACTC CTGGCTTAAG TAGTTTAGT- 960

CCTGGGAGAA TCATTAAGTG AAAAGCATTI AACTTAAAA AAAAAAAAAA AAAAAAAAAA 1020

AAACSTGGTG CGGAATTGG CACGAGC1AA CCGAGAAACA TCCAATTCTC AAACCTGAAGC 1080

TCGCACTCTC GCGTCCAGCA TGAAAGTCTC TGCGGCCCTT CTGTGCCTGC TGCTCATAGC 1140

AGGCACCTTG ATTCCCCAAG GGCTGGCTCA GGCAGATGCA ATCAATGCCC CAGTCACCTG 1200

CTGYTATAAC TTACCAATA GGAAGATCTC AGTCAGAGG CTCGCGAGCT ATAGAAGAAC 1260

CACCAGCAGC AAGTGTCCCA AAGAAGCTGT GATCTTCAAG ACCATTGTGG CCAAGGAGAT 1320

60

CTGTGCTGAC CCCAAGCAGA AGTGGGTTC A GATTCATG GACCACCTGG ACAAGCAAAC 1380
 CCAAACTCCG AAGACTTGAA CACTCACTCC ACAACCCAAG AATCTGCAGC TAACTTATTT 1440
 5 TCCCCAGCT TCCCCAGAC AACTCTTTTT ATTTTATTAT AATGAATTTT GTTGTGTGAT 1500
 GTGAAACATT ATGCCTTAAG TAACTTTAAT TCTTATTTAA GTTATGTATG TTTTAAGTTT 1560
 ATCTTTCATG GTACTAGTGT TTTTAGATA CAGAGACTTG GGGAAATGCG TTTTCTCTTT 1620
 10 GAACCACAGT TCTACCCCTG GATCTTTTTG AGGGTCTTTG CAAGAATCAT TAATACAAAG 1680
 AATTTTTTTT AACATTCCAA TGGATTGCTA AAATATTATT GTGGAAATGA ATATTTTGTA 1740
 15 ACTATTACAC CAAATAAATA TATTTTGTGA CAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1800
 AAGSGGCCGC TCGAATTAAG CC 1822

20

(2) INFORMATION FOR SEQ ID NO: 106:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1712 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:

CGTGCCCCAG CCTCCCGAGT AGCTGGRACT ACAGGCACGT SCCACCACGC CCAGCTAATT 60
 TTWATATTTT WAGTAGAGAC GGGGTTTTSC TGTKTGGCC AGGCTGGTCT CAAACTCCTG 120
 35 ACCTCAAGTA ATCCACCTGG CCTGTCTTTT TCATGTCTTA ACATGGCATG TCTTTTAGTT 180
 TCATTATTTT CCTACTCCTT GTATGTCAAG AAATTACATT TTGCATGTCT TATGGAGATG 240
 40 CTGTTAATTG CTTCAGTGAG TGCTTTTCTA ATCTGCAGAC CATTTACATT TCCTGTTTGC 300
 AGCATGCTGT GTGCAAACAC TCAGTAATTT GGAGTATTCA ATTATTTGTT AGGGCTCTTC 360
 CTATTTCCAA ATGTGCTGAA TTGTCTATTG ATGGGATTTT CAGATCTTTT CATGAGAACT 420
 45 GGAAATGTAG CTGGGTGGCA CCTACCTAGG TTGCTACGTA GTGAGTAGAC TTTCTCTTGG 480
 GTATAGTAAG CCTCAGACAG CTTTCACTTT TATCTACTTT ACTTGTGGAA ATAAAACAGT 540
 50 CATTTTGTTC TGAAAGAATA AGATAGCTTT CTGTAGAGAA GGAATTCCTA CCTCTAAAAG 600
 CTGCCTTGAG AACTCAGAAC TGGCAGTTTT CTGAGGTGAT TTTTAAATTT CAGTATTAGG 660
 GAGAGTCCAG CATTTGCTGA CACAGATTCT ACATAACTAA TGTATGATAG CAAATGCAAA 720
 55 ACTATTATAA TGTGGTGTAT CTGTGCATA CACAGGTTAG AACAAAGTAG CTCTGGCAGT 780
 AGATCTCCAG AGACCCAAGT TTAGTTCTCT ATAGTGTATT TGAAGTAGTT ATACTCTCTG 840
 60 CTTAAGTAGT TTAGTGCCTG GGAGAATCCA TTAAGTAAAA GCATTTAACT TAAAAAATAA 900

AAAAAAAAAA AAAAAAAAAA CTCTGCGGA ATTGGGACG AGCAGAAACA TCCAATCTC 960
 AACTGAAAG TGGCACTCTG GCTTCAGCA TGAAGTCTC TGGGCCCCCT CTGTGCTGC 1020
 5 TGCTCATAG AGCCACCTTC ATTGCTCAAG GCTGGCTCA GGCAGATGCA ATCAATGCC 1080
 CAGTCACCTG CTGTATTAAG TTAACAATA GGAAGATCTC AGTGAGAGG CTGGGAGCT 1140
 10 ATAGAAGAA CACCAGCAGC AATGTCCCA AAGAAGCTGT GATTTCAAG ACCATTGTG 1200
 CCAAGGAGAT CTGTGCTGAC CCAAGCAGA AGTGGGTCA GATTCATG GACCACCTG 1260
 ACAAGCAAAC CCAACTCCG AAGATTBAA CACTCACTCC ACAACCCAG AATCTGCAGC 1320
 15 TAACTIATTT TCCCCTAGCT TTCCCGAGC ACCCTGTTTT ATTATTTAT AATGAATTT 1380
 GTTGTGAT GTGAAACAT ATGCTTAAG TAATGTTAAT TCTIATTTAA GTIATTGATG 1440
 20 TTTTAAGTT ATCTTTCATG GTACTAGTGT TTTTAGATA CAGAGACTTG GGGAAATGC 1500
 TTTTCTCTT GAACCACAGT TCAACCTG GATGTTTTG AGGGTCTTG CAAGAATCAT 1560
 TAATACAAAG AATTTTTTTT AACATTCCAA TGCATTGCTA AATATTATT GTGGAAATGA 1620
 25 ATATTTTGTA ACTATTACAC CAAATAAATA TATTTTGTA CAAAAAAAAA AAAAAAAAAA 1680
 AAAAAAAAAA AAGSGCCGC TCGAATTAAG CC 1712

30

(2) INFORMATION FOR SEQ ID NO: 107:

35

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1969 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 107:

CCCCTCCTTC CCCTGCCCAC CTACTGAACC CTCCTCCGAG GTGCCCGAGC AGCCGTCTGC 60
 45 CCAGCCACTC CCTGGGAGTC CCCCAGAAG AGCCTATTAC ATCTACTCCG GGGGCGAGAA 120
 GATCCCCCTG GTGTTGAGCC GGCCCTCTC CTCCAACGTG GCCACTCTTC AGCATCTCTG 180
 TCGGAAGACC GTCAACGGCC ACCTGGACTC CTATGAGAAA GTCACCCAGC TGCCGGGGCC 240
 50 CATTCGGRAG TTCTGGACC AGTACGATGC CCCGTTTTAA GGGGTAAAGG GCGCAAAGGG 300
 CATGGGTCGG GAGAGGGGAC GCAGGCCCCCT CTCCTCCGTG GCACATGGCA CAAGCACAAG 360
 55 AAGCCAACCA GGAGAGAGTC CTGTAGCTCT GGGGGGAAAG AGGGCGGACA GGCCCTCCC 420
 TCTGCCCTCT CCTGCAGAA TGTGGCAGGC GGACCTGGAA TGTGTGGAG GGAAGGGGGA 480
 GTACCACCTG AGTCTCCAGC TTCTCCGGAG ACCCAGCTGT CTTGGTGGGA CGATAGCAAC 540
 60

CACAAGTGGG TTCTCCTTCA ATTCTCTCAGC TTCCCTCTCTG CCTCCAAACA GGGGACACTT 600
 CGGGAAATGCT GAATTAATGA GAACTGCCAG GGAATCTTCA AACTTTCCAA CGGAACCTGT 660
 5 TTGCTCTTTG ATTGCTTTA AACCTGAGCT GGTGTGGAG CCTGCGAAAG CTGGAAGAGA 720
 GAGAGGTCTT GAGGCCCCCA GGGGTGGGG CTGGCGAAGG AAATGCTCAC ACCCCCCGCC 780
 CACCCAGGC GAGGATCTG GTGACATGCT CCTCTCCCTG GCTCTGGGGA GAAGGGCTTG 840
 10 GGGTGACCTG AAGGGAACCA TCCTGCTGCC CCACATCTCT TCCTCCGGGN ACAGTCACCG 900
 AAAACACAGG TTCCAAAGTC TACCTGGTGC CTGAGAGCCC AGGGCCCTTC CTCCGTTTTA 960
 15 AGGGGGAAGC AACATTTGGA GGGGACGGAT GGGCTGGTCA GCTGGTCTCC TTTCTCTACT 1020
 CATACTATAC CTTCCTGTAC CTGGGTGGAT GGAGCGGGAG GATGGAGGAG ACGGGACATC 1080
 TTTCACCTCA GGCTCCTGGT AGAGAAGACA GGGGATTCTA CTCTGTGCCT CCTGACTATG 1140
 20 TCTGGCTAAG AGATTCTGCT TAAATGCTCC CTGTCCCATG GAGAGGGACC CAGCATAGGA 1200
 AAGCCACATA CTCAGCCTGS ATGGGTGGAG AGGCTGAGGG ACTCACTGGA GGGCACCAAG 1260
 25 CCAGCCCA CA GCCAGGGAAG TGGGGAGGGG GGGCGGAAAC CCATGCCTCC CAGCTGAGCA 1320
 CTGGGAATGT CAGCCAGTA AGTATTGGCC AGTCAGGCGC CTCGTGGTCA GAGCAGAGCC 1380
 ACCAGGTCCC ACTGCCCGA GCCCTGCACA GCCCTCCCTC CTGCCTGGGT GGGGGAGGCT 1440
 30 GGAGGTCATT GGAGAGGCTG GACTGCTGCC ACCCCGGGTG CTCCCGCTCT GCCATAGCAC 1500
 TGATCACTGA CAATTTACAG GAATGTAGCA GCGATGGAAT TACCTGGAAC ATTTTTTGT 1560
 35 TTGTGTTTG TTTTGTGTTT TGTGGGGGGG GGCAACTAAA CAAACACAAA GTATTCTGTG 1620
 TCAGGTATTG GGCTGGACAG GGCAGTTGTG TGTGGGGTG GTTTTTTCT CTATTTTTTT 1680
 GTTGTGTTCT TGTGTTTTAA TAATGTTTAC AATCTGCCTC AATCACTCTG TCTTTTATAA 1740
 40 AGATTCACCC TCCAGTCTCT TCTCTCCCC CCTACTCAGG CCCTTGAGGC TATTAGGAGA 1800
 TGCTTGAAGA ACTCAACAAA ATCCCAATCC AAGTCAAACT TGGCAGATAT TTATATTTAT 1860
 45 ATTGAGAAAA GAAACATTTT AGTAATTTAT AATAAGAGC ACTATTTTTT AATGAAAAAA 1920
 AAAAAAAAAA AAAAAAAAAA CGACGCTGGT GACCGGAATY CGACGTACG 1960

50

(2) INFORMATION FOR SEQ ID NO: 108:

55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1734 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:

(2) INFORMATION FOR SEQ ID NO: 109:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2003 base pairs

(E) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 109:

CGCAGGGGGG GGGGGGGGGG GGGAGTCGCA TTCCCCCGGT CCCCCCTCCAC CCCACGGGGG 60
 15 CTGGACCATG GATGTCAGAT GGTGGGCACT GSTGTGTGTG GTTGGCTTCC CTTCCCTAGG 120
 GGCAGGTGGG GAGACTCCCG AAGCCCCCTC GAGTCATGG ACCCAGCTAT GTTCTTCCG 180
 20 ATTTGTGGTG AATGCTGGTG GTTATGCCAG NNTTATGGA CTTGGCTACC TCTGGTGCA 240
 CTACTTCAGG CGAAGAAAT ACCTGGAGAC CGTAGGGGC CTCTGTTTC CCGTGGTGA 300
 AGCTGTGTG TTGAGCAATG AGCCCCAGGC CTCTGATGAG GTTCCCCCG CGCCCCGAA 360
 25 AGAGGGCGCA GAGACACCC CATGTGGCA GGCCTGAAG CTGCTCTTCT GTGCCACAGG 420
 GCTCCAGGTG TCTTATCTGA CTTGGGTGT GTTGCAAGAA AGATGATGA CCGCAGCTA 480
 TGGGGCCACA GTCACATCAC CGGGTAGCG CTTACGAC TGGCAGTCC TGGTGCTAA 540
 30 GAACCGAGTG CTGACTCTGA TTGTGGCTGG CCTCTCTGT GTTCTCTGCA AGCAGCCCCG 600
 GCATGGGGCA CCACTGTACC GGTACTCCTT TGCCAGCCTG TTTAATGTGC TTAGCAGCTG 660
 35 GTGCCAATAC GAAGTCTTA AGTTCGTGAG CTCCCCACC CAGGTGCTGG CCAAGGCCTG 720
 TAAGGTGATC CTTGTCATGC TGATGGGAAA GTTGTGTCT CCGCGCANTA ACGAACAATG 780
 GGAGTACCTG ACAGTCACCC TCATCTCCAT TGGGGTCAGC ATGTTTCTGC TATCCAGCGG 840
 40 ACCAGAGCCC CGCAGCTCCC CAGCCACCAC ACTCTCAGGC CTCATCTTAC TGGCAGGTTA 900
 TATTGCTTTT GACAGCTTCA CTTCAAATG GCAGGATGCC TTTTGCCTA TAAGATGTCA 960
 45 TCGGTGCAGA TGATGTTTGG GSTCAATTTT TTCTCTGCC TCTTACAGT GGGSTCACTG 1020
 CTAGNAACAG GGGGEMCCTA CTGGAGGGAA CCGCTTCAT GGGGCGACAC AGTGAGTTTG 1080
 CTGCCCATGC CTTGTTACTC TCCATCTGCT CGCATGTGG CAGCTCTTC ATCTTTTACA 1140
 50 CCATGGGGCA GTTGGGGGT GCGTCTTCA CCATCATCAT GACCTCCGC CAGGCTTTG 1200
 CCATCTTCT TCTGCTCTT CTCTATGGC AACTGTAC TGTGGTGGG GGGCTGGGG 1260
 55 TGGCTGTGGT CTTGCTGCC CTCTGTGCA GAGTCTAGC GCGGGCCGT CTAAAGCAAC 1320
 GGGGAAAGAA GCTGTGCCT GTTGTGCTC CTGTGCAGAA GCTTTGAGG TGGAAAGGG 1380
 60 CTGAGGGGTG AAGTGAAATA GGACCTCCC ACCATCCCCT TCTGCTGTAA CCTCTGAGG 1440

AGCTGGCTGA AAGGGGAAAA TGCAGGTGTT TTTCTCATTAT CACAGACCAG CTCTGAGCA 1500
 GGGGATTGAG GAGCCAGGA GGCAGCCTTC CTTTTCCT TAAGTCACCC ATCTTCAGT 1560
 5 AAGCAGTTTA TTGTGAGCCC CGGGGTABA CACTCTCAG TGAGGGGTTT TGGGAGTTT 1620
 GGGGTCAAGA GAGCATAGGT AGGTTCACA GTTATTCTTC CCACAAGTTC CTTAAGTCT 1680
 TGCCCTAGCT CTCTCTGCC ACCTTCAGA CTCACCTCCC TGTCAAATA CTTGCATTTC 1740
 10 TTACCTGCT GAGAAAAGCA CAAGCGTGT AGGTCCAAT GCTGCTTCC CAGGAGGGTG 1800
 AAGATGCTGC TGTGCTGAGG AAAGGGGATG CAGACCCCTG CCCAGCACCA CCACCTCCTA 1860
 15 TGCTCTGGA TCCCTAGGCT CTGTTCATG AGCTGTTC AGGTTTGGT ACTTTAGAAA 1920
 TGTAACTTTT TGTCTTATA ATTTATTTT ATTAAATTAA ATTACTGCAA AAAAAAAAAA 1980
 AAAAAAATCG GGGGGGGGCC CGN 2003
 20

25 (2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1320 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 30 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 110:

GCTGAGCTGC CTTGAGGTGC AGTGTGGGG ATCCAGAGCC ATGTGGGACC TGCTACTACT 60
 35 GGGCCTGATT GGGGGCTGA CTCTCTACT GTTGTGACG CTGCTGGCCT TTGCCGGTA 120
 CTCAGGCTA CTGGCTGGG TGAAGTGAG TCTGGGTCA CCCCCATCC GCAACGTAC 180
 40 TGTGGCCTAC AAGTTCCACA TGGGGCTCTA TGGTGAGACT GGGCGGCTTT TCACTGAGAG 240
 CTGCAGCATC TCTCCAAGC TCCGTCCAT CGCTGTCTAC TATGACAACC CCCACATGCT 300
 GCCCCCTGAT AAGTGCCGAT GTGCCGTGGG CAGCATCTG AGTGAAGGTG AGGAATCGCC 360
 45 CTCCCCGAG CTCATCGACC TCTACCAGAA ATTTGGCTTC AAGGTGTTCT CTTCCCGGC 420
 ACCCAGGAT GTGGTGACAG CCACCTTCCC CTACACCACC ATTCTGTCCA TCTGGCTGGT 480
 50 TACCCGCTGT CTCATCTCTG CCTTGACAC CTACATCAAG GAGCGGAAGC TGTGTGCTA 540
 TCCTCGCTG GAGATCTACC AGGAAGACCA GATCCTATTC ATGTGCCAC TGGCAGGCA 600
 GGSAGACTTC TATGTGCTG AGATGAAGGA GACAGAGTG AAATGGCGGG GCTTGTGGA 660
 55 GGCCATTGAC ACCCAGGTGG ATGGCACAGG AGCTGACACA ATGAGTGACA CGAGTTCTGT 720
 AAGCTTGGA GTGAGCCCTG GCAGCCGGGA GACTTCAGCT GCCACACTGT CACCTGGGG 780
 60 GAGCAGCCGT GGCTGGGATG ACGGTGACAC CCGTAGCGAG CACAGCTACA GCGATCAGG 840

255

5 TGCCAGCGGC TCCTCTTTTG AGGAGCTGGA YTTGGAGGGC GAGGGGCCCT TAGGGGAGTC 900
ACGGCTGGAC CCTGGGACTK AGCCCCCTGGG GACTACCAAG TGGCTCTGGG AGCCCACTGC 960
CCCTGAGAAG GGCAAGGAGT AACCCATGGC CTGCACCCTC CCTGCAGTGC AGTTGCTGAG 1020
GAACTGAGCA GACTCTCCAG CAGACTCTCC AGCCCTCTTC CTCCTTCCTC TGGGGGAGGA 1080
10 GGGGTTCCCTG AGGGACCTGA CTCCCCCTGC TCCAGGCCTC TTGCTAAGCC TTCTCTCAC 1140
TGCCCTTTAG GCTCCCAGGG CCAGAGGAGC CAGGGACTAT TTTCTGCAAC CAGCCCCCAG 1200
GGCTGCCNCC CCTGTTGTGT CTTTTTTTCA GACTCACAGT GGAGCTTCCA GGACCCAGAA 1260
15 TAAAGCCAAT GATTACTTG TTTCAAAAAA AAAAWAAAAA AAAAAAAAAA AAAAAAAAAA 1320

20

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:
25 (A) LENGTH: 1962 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:

CGGACCCCTT CCTCTCCTC NAAGCATGTC CCACCATTGT GGCAGGGGCT GGGGANACAG 60
TCACCTGATG CGGGGACCAC GGCCACTCCA CCTCGSTGGC GCTGTCACTG GGCACTACTG 120
35 GCTGGGCCTG CACTGAGGTC CCTGCTGGGG CAGTTCTTCC AGAATTATCT TCAGAGGGGG 180
CCTCCAGCTC CCTGGTACCC TCAGGGGCCC GTGTGGCTGG AAGCAGGGAA GGGGCACCCT 240
CGGAGCTTCC TGTCTCCTCG CTCTCTCCTC GAGGGACCCC AGATAGCTCA GGACCACCAG 300
40 TTGCTCCCC CACCTCTCTT GCCTCAACCA GAGTGAAGG TGATGGGGAT GCTAGGTTCC 360
TCTCCCTGGG AGTGGGCAGA GTCTCAGTAG CTGTCCATG GACCCCTGGA GGCTGGAAG 420
45 CTTCTGACTC TCCATCAGGA AGTGGTGATG CACCAGGCTG CAGGACTGCC CTTGCTGGCG 480
CCTGGGAGAG TGACTCCTCC TGGGCTGCTG GCTCAGTGGG GAGAGAGGCC TCAGGGCCCC 540
GGCTGCTGAG CTCGCTGGGC CATGCCCCA GAGCCTCATC CTCCACCTCC TCCTCTCTCT 600
50 CTTCTCTCTC TTTCTCTTCT TCATCTTCAT ATTTCTCTTC TTCTCCAAT GCCTTACCTT 660
CCTCTTYTGR AAACCCCGTG GCGGTACCA TGGATTGTGT TTCAAATTCT AGGAGCGTCC 720
55 TAGGGGCCTC TGCTGGGTCT TCTGGAGTGG AGCTTCCACC TCCTCCGTCC TCCATGATGG 780
GGATGGAGTA RATGGCCCCA CGGGATTAC TCTCTGTGGC TTCTGAGGC AGCTGCAGTI 840
CCTCCAGGGT CTCTGTCACT GTGACRATAG CCTCTAGTCC ATCAAAGCT GGSTTGAGG 900
60

CTGGGTTGGA GGGCTCAGGG ATGGAAG GCTGGGCCGA GTCTGGGAAG CAGTARACGT 960
 TGAAGGGGCT GTGTTTATTG GGAAGTCAAG TCTGGTTGGG GAAGANGAAG AGAGTCTTGA 1020
 5 CACCAGGCAA GCTCCACCA CAGGCTKGD TGGTGTGAC GATGGGTAG CGCACANTGC 1080
 CATCAGTAG CCACTGGGC TGCATGCTC CAGGCCACCA TCCAGGCTG CATACAGTTG 1140
 GCGGTGTG GAAATCTCTG CACCCGCTG CTGGCAGTAC GTCCTGCTT COTCCAATGT 1200
 10 CAGCTTCTCT GGAGGCTCAG CCAGGAACAG TTCTCCATTT AGGTCTTCTAG CATAACAGTA 1260
 CACATCATAG AGGTCATCCG GGTCCACCAC ACCATAGTTC CGGACCCCGG GGAAGCCATC 1320
 15 CATGTCTCCG TAACAGGGCT CTGCTGGGTG CTGSATGGGA TACCTTTGAC CTTGAMCTCC 1380
 ACAGCTGCGC TGCTGTGATC GATGGCTGC TGGACCTCAC AGCGATAGAT ACCTGAGTCG 1440
 TTGGGGCGCA GCTCGCTCAG CGTCAGGGA GACGTGGGTG AGCGACGCTG GGTACGCAGG 1500
 20 CAGTGCCACG CGGAACGGT AGGCTCTCTT CACCTTGACG CGCACTCCCC GCGCCACCAG 1560
 CACTCTGCC TCCCGCCCC GGAACAGGAA AGTCCACTTG ACCCGCGGAG AGCCACAGCAG 1620
 25 AGCCCGCGCG CTCGGCGGTG SCCGAGGTA GTGGACGTGG CAAGGGATGK TGAGGGCSCC 1680
 GCCGAGCAAC GCCYTGCAGT GGGCGCTGCG CCGCGATGCG CACCGGAAA GCGCGKTCCT 1740
 CTGAGCTGTC TCCTTCCAGA ACATCTGCTA AAGCTGCAGG AGCCTGGGCC AGGACCAGGG 1800
 30 CTGCCAGCAG GGCAGGAAC AGCTGGGCCA TGCTGCAGGC TACCCAGGGC TGGGGTTGGG 1860
 TCGCGGCACT GCGAAGTTTG TCGCTCTCTC CGGGGTCTC CTCGGGCTKC ACGGCTCAGT 1920
 35 NCCTGCAGCT GCAGCTGAGA CTGCGGCGGA GACTGCGCGA GC 1962

40 (2) INFORMATION FOR SEQ ID NO: 112:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1785 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:

50 AAGTTTCAGC CAAACTTCGG GCGGCTGAGG CGGCGGCCGA GGAGCGGCGG ACTCSGGGCG 60
 CGGGGAGTCG AGGCATTGTC GCCTGGGCTT CGGAGCGTAC CGCAGGGCCT GAGCCTTTGA 120
 AGCAGGAGGA GGGGAGGAGA GAGTGGGGCT CCTCTATCGG GACCCCTCC CCATGTGGAT 180
 55 CTGCCCAGGC GCGCGGCGCG GCCGAGGAGG CGACCGAGAA GATRCCCGCC CTGCGCCCCG 240
 CTCTGCTGTG GGGCTGCTG GCGCTCTGGC TGTGCTGCGC GACCCCGCGC ATGCATTGCA 300

GTGTGAGAT GGCTATGAAC CCTGTGTAAA TGAAGGAATG TGTGTTACCT ACCACAATGG 360
 CACAGSATA TGC AAATGTC CAGAAGGCTT CTGCGGGGAA TATTGTCAAC ATCGAGACCC 420
 5 CTGTGAGAAG AACCGCTGCC AGAATGCTGG GACTTGTGTG GCCDAGGCCA TGCTGGGGAA 480
 AGCCACGTGC CGATGTGCCT CAGGCTTTAC AGSAGAGGAC TGCCAGTACT CGACATCTCA 540
 10 TCCATGCTTT GTGTCTCGAC CTGCTCTGAA TGGCGGCACA TGCCATATGC TCAGCCGGGA 600
 TACCTATGAG TGCACCTGTC AACTCGGGTT TACAGGTAAG GAGTGCCAAT GGACCGATGC 660
 CTGCCTGTCT CATCCCTGTS CAAATGGAAG TACCTGTACC ACTGTGGCCA ACCAGTTCTC 720
 15 CTGCAAATGC CTCACAGGCT TCACAGGGCA GAAGTGTGAG ACTGATGTCA ATGAGTGTGA 780
 CATTCCAGGA CACTGCCAGC ATGGTGGCAC CTGCCTCAAC CTGCCTGGTT CCTACCACTG 840
 CCAGTGCCTT CAGGGCTTCA CAGGCCAGTA CTGTGACAGC CTGTATGTGC CCTGTGCACC 900
 20 CTCGCCTTGT GTCAATGGAG GCANCTGTG GCAGACTGGT GACTTCACTT TTGAGTGCAA 960
 CTGCCTTCCA GAAACAGTGA GAAGAGGAAC AGAGCTCTGG GAAAGAGACA GGAAGTCTG 1020
 25 GAATGGAAAA GAACACGATG AGAATTAGAC ACTGGAAAAT ATGTATGTGT GGTAAATAAA 1080
 GTGCTTTAAA CTGAATTGAC ATTAAACAGT GGTGATCAAC TTTMCTATGT GCTTGTGCTT 1140
 TTGCTTTTGA TGGAGTAATT CATTGTTTTT TTATCCACCT AAATGCACCC AGCTGCCCTT 1200
 30 GATTTTCTCT GGGCTACTGG CCTTCACAAC CCTCTCCCAT GTACCTCTCTC TGACTTTGGG 1260
 GTAACCTCC CTAACCTTAA AGCTAGAGAA TTCTGAAACT GAGGAGGGGA TCCTCTGTTA 1320
 35 ATCACTGAGC ACTTTTTGAT GAGCTGATAG ATGATATATG AGAGACTATG CGTGGCACAA 1380
 TACTTTGTTA CACTCTTCAC TGATACAAGT GTTCTAGAGT GYACACACAA CCCAAAGATA 1440
 GAAATAAAAA GAGGAGCAGT GTCGGGGAGC TTGGGGCCTG GTGTTCCATG GAGAGGGAGA 1500
 40 AAGGAACAAG CTGRCCAAT TCATTCAACT CCTTATAAAA ATGATGAGGA GGCTGAAAAC 1560
 CAAGAATTTT GATTGGGAAC AGAATACAAG CAGCTGAAC AGATGAWTTA CTAAGCAACA 1620
 45 AAGATCCIGT TTTTATACAA ATATCCTTAG TACAAAAACA AAARAAGGAA AACTGTAGGG 1680
 GGGAGTAATG TGCTAAGTAA GCAGAAATGC CTCCAAAAGA AGTTGTTTCT AGTTACTCTT 1740
 50 TTCCGGGTNG GGATCTTTAG NTCCGGTAT TGTGGGTATG GTTCC 1780

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1842 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

5	GGAGCCCTCTC TTGCAACTTC TGACACCGCG GGCACCGCG GCGGCTGAT CCGGCAGAGG	60
	AAGTCGCGGG CCTGAGACGA TGACCGCGCG GGTTCGCGG GAGCGCGCG GGTTCGCACA	120
	GCGCGCGCG CTCTGCTGCG TGTGCTGCT GCGCTGCTG CTAGTACCG CGGAGCGCG	180
10	GAAACCTGCA GGAGTCTACT ATGCAACTGC ATACTGATG CCGCTGAAA AGACAGTACA	240
	AGTCAAAAAT GTATGACGA AGAATGGGGA CGCTATGCG TTTCACAATA ACTCTGTGAA	300
	AACCACAGGG TGGGCGATCG TGAGATCAG AGCTGCTAT GGTCTGAAA CCGTGAGCAA	360
15	TGAGATCATC ATGTTTGCGS CTGCTTTTTT GGAGGCTTAC CTCACGCGC CACACATGAA	420
	TSACCACTAC ACAAACCTCT ACCACAGGT GATCAGGAAA CCGTTCATCA TGGATAAAGT	480
20	GCAGGATTTT ATGAGAGAAGT AAGATAAGTG GACCGGAAA AATATCAAAG AATACAAGAC	540
	TGATTCATTT TGAGACATA CAGGTATGT GATGACAAA ATAGATGCGC TCTATGTAGG	600
	AGCAAAGAAG AGGCTATAT TAGAAGGAC AAAGTCAATG ACCCTGTTCC AGATTGAGTT	660
25	CCTGAATAGT GTTGAGATC TATTGATCT GATTCCTCA CTCTCTCCA CAAAAACGG	720
	CAGCCTAAAG GTTTTTAAGA GATGAGACAT GGCACATTG TCGCTCTTA TCAAGGTTCT	780
30	TCCTGGATTT GAGAACATCC CTTTGTCTCA CTCAGCTGG TACACGTATG CAGCCATGCT	840
	CAGGATATAT AAACACTGGG ACTTCAACRT CATAGATAAA GATACCAGCA GTAGTCGCCT	900
	CTCTTTCAGC AGTTACCCAG GGTTTTGGG GTCTCTGGAT GATTTTTACA TTCTTAGCAG	960
35	TGGATTGATA TTGCTGCGA CCACAAACAG TGTGTTTAA AAAACCTGC TAAAGCAGTA	1020
	ATACCCGAGA CTCTCTGTC CTGGCAAAGA GTCCGTGTG CCAATATGAT GGCAGATAGT	1080
40	GGCAAGAGGT GGACAGACAT CTTTCAAAA TACAACTCTG GCACCTATAA CAATCAATAC	1140
	ATGGTTCTG ACCTGAAGAA AGTAAAGCTG AACCACAGTC TTGACAAAG CACTCTGTAC	1200
	ATTGTGGAGC AAATTCCTAC ATATGTAGAA TATTCTGAA AACTGATGT TCTACGGAAA	1260
45	GGATATTGGC CCTCTACAA TGTTCTTTC CATGAAAAA TCTACAACTG GAGTGGCTAT	1320
	CCACTGTTAG TTCASAAGCT GGGCTTGGC TACTTTIATG ATTTAGCTCC ACGAGCCAAA	1380
50	ATTTTCCGGC GTGACCAAGG GAAAGTGAAT GATACGGCAT CCAATGAAAT TATCATGCGA	1440
	TACAACAATT ATAASAAGGA TCCTTACAGT AGAGGTGACC CCGTAATAC CATCTGCTGC	1500
	CGTGAGGACC TGAATCACC TAACCAAGT CCGGAGGTT GTTATGACAC AAAGGTGGCA	1560
55	GATATCTACC TAGCATCTCA GTACACATCC TATGCCATAA GTGGTCCAC AGTACAAGGT	1620
	GGCTCCCTG TTTTTCGCTG GGACCGTTC AACAAAATC TACATCAGG CATGSCAGAG	1680
60	GTCTACAACT TTGATTTTAT TACCATGAAA CCAATTTTGA AACTTGATAT AAAATGAAGG	1740

AGGGAGATGA CGGACTAGAA GACTGTAAAT AAGATACCAA AGGCACTATT TTAGCTATGT 1800
 TTTTCCCATC AGAATTATGC AATAAAATAT ATTAAATTGT CA 1842

5

(2) INFORMATION FOR SEQ ID NO: 114:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1960 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 114:

GAATTGCGCA CGAGTTCTCT CGCGCCCCAG CGGCGGCTG CCAGCTTTTC GGGGCCCCGA 60
 GTCCGACCCA GCGAAGAGAG CGGCCCCGGG ACAAGCTCGA ACTCCGGCCG CCTCGCCCTT 120
 CCCCAGCTCC GCTCCCTCTG CCCCCTCGGG GTGCGCGCC CACGATGCTG CAGGGCCCTG 180
 GCTCGCTGCT GCTGCTCTTC CTCGCCTCGC ACTGCTGCCT GGGCTCGGCG CGCGGGCTCT 240
 TCCTCTTTGG CCAGCCCGAC TTCTCCTACA AGCGGAGMAA TTGCAAGCCC ATCCCGGTCA 300
 ACCTGCAGGT GTGTCACGGC ATCGAATACC ABAACATGCG GTTGCCCAAC CTGCTGGGCC 360
 ACGAGACCAT GAAGGAGGTG CTGGAGCAGG CCGGCGCTTG GATCCCGCTG GTCATGAAGT 420
 AGTGCCACCC GGACACCAAG AAGTTCTGT GTGCTCTCTT CGCCCCCGTC TGCCTCGATG 480
 ACCTAGACGA GACCATCCAG CCATGCCACT CGCTCTGCGT GDAGGTGAAG GACCGCTGCC 540
 CCCCAGTCAT GTGCGCCTTC GGNITCCCCCT GGGCGGACAT GCTTGAGTGC GACCGTTTCC 600
 CCCAGACAA CGACCTTTGC ATCCCCCTCG CTAGCAGCGA CCACCTCCTG CCAGCCACCG 660
 AGGAAGCTCC AAAAGTATGT GAAGCCTGCA AAAATAAAAA TGATGATGAC AACGACATAA 720
 TGGAAACGCT TTGTAATAAT GATTTTGCAC TGAAAATAAA AGTGAAGGAG ATAACCTACA 780
 TCAACCGAGA TACCAAAATC ATCCTGGAGA CCAAGAGCAA GACCATTAC AAGCTGAACG 840
 GTGTGTCCGA AAGGACCTG AAGAAATCGG TGCTGTGGCT CAAAGACAGC TTGCAGTGCA 900
 CCTGTGAGGA GATGAACGAC ATCAACGCGC CCTATCTGCT CATGGGACAG AAACAGGGTC 960
 GGGAGCTGCT GATCACCTCG GTGAAGCGGT GGCAAGAGGG GCAGAGAGAG TTCAAGCGCA 1020
 TCTCCCGCAG CATCCGCAAG CTGCAGTGCT AGTCCCGGCA TCCTGATGGC TCCGACAGGC 1080
 CTGCTCCAGA GCACGGCTGA CCATTCTGTC TCCGGATCT CAGCTCCCGT TCCCCAAGCA 1140
 CACTCCTAGC TGCTCCAGTC TCAGCCTGGG CAGCTTCCCC CTGCCTTTTG CACGTTTGCA 1200
 TCCCCAGCAT TTCCTGAGTT ATAAGGCCAC AGGAGTGGAT AGCTGTTTTT ACCTAAAGGA 1260

60

AAAGGTCAGC CGAATCTTGT AGAATATATC AAAGTAATAA AATCATGAAT ATTTTATGA 1320
 AATTTAAAAA TAGCTTACTT TAAAGTATG TTGAATAGG TCGAAGTGTG ACTTGAGTGT 1380
 5 GGTTCGTTGT TGTTCGTTGT TTGAGTTCAG CTGATTTTCA CTTCGCACTG AGTTTTCAT 1440
 AACATGCAAA TCGTTCAAT TTCTCTGTC GCGCAAACTT GTGGTTCACA AAGGCTGTTG 1500
 AGATAAAGCT GCGTGTATG TCAACATCTT CATCAGCTCC AGACTGAGAC TCACTTCTTA 1560
 10 AGTCTTACAA CAATTCATCA TTTTATACCT TCAATGGGAA CTAAACTGT TACATCTATC 1620
 ACATTCCAGC TACAATACTT CCATTTATTA GAAGCACATT AACCATTTCT ATAGCATGAT 1680
 15 TTCTTCAAST AAAAGGCCAA AGATATAAAT TTTATAATTG ACTTGAGTAC TTAAAGCCTT 1740
 GTTAAAAACA TTCTTACTT AACTTTTGCA AATTAAACCC ATTGTAGCTT AACTGTAATA 1800
 TACATAGTAG TTTACCTTTA AAAGTTGTA AATATTGCT TTAACCAACA CTGTAATAT 1860
 20 TTCAGATAAA CATATATTC TTCTATATAA ACTTACATC CTGTTTACC TAAAAAAA 1920
 AAAAAAAAAA AAAAACTCG AGGGGGGCCC GGTACCCAAT 1960
 25

(2) INFORMATION FOR SEQ ID NO: 115:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 536 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GTGCTCAGCC GCGGGGGCAC AGYAGGACGT TTGGGGGCGT TCTTTCAGCA GGGGACAGCC 60
 40 CGATTGGGGA CAATGGCGTC TCTTGGCCAC ATCTTGTTT TCTGTGTGG TCTCCTCAGC 120
 ATGGCCAAAG CAGAAAGTCC AAAGGAACAC GATCCGTTCA CTTACGACTA CCAGTCCCTG 180
 CAGATCGTAG GCTTCGTCAT CGCGGGGATC CTCTTCATCC TGGGCATCCT CATCTGCTG 240
 45 AGCAGAGAT GCGGTGCA GTTCAACCA CACGAGAGGA CTGGGGAACC CGATTAAGAG 300
 GAGGAACTT TCGGAGGTC CATCCGCTCT CTCTCAGC GCATCCGCTA GAAAGACTG 360
 50 GAGGATGGA ATCGGGCCAG GACTCCCTG GCACCTGATA TCTCCACGC TCTACCTGCG 420
 CGCCACATG CCGTTCGCG GCGCTTCCG CAGCCCTGCC CCGGAGACT CCGCTGCGG 480
 CCAAGACTG CATATAAAG TCGTTCTCT TCGAAAAA AAAAAATAA AAAACT 536
 55

60 (2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 790 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 116:

GTGGGGAGGG GCGGGAGCAA AGCCGCGCCT CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC 60
 10 CTGACTTGAA CCTTCCCGGT CCCAGCCCT CAACAGGAGG CGCAGAAAAT CTTCAAAGCC 120
 AACCACCCCA TGGACGCAGA AGTTACTAAG GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC 180
 15 CTGGACAATG TGGACCCCAA CCTGAGAAC TTCGTGGGGG CGGGGATCAT CCAGACTAAA 240
 GCCCTGCAGG TGGGCTGTCT GCTTCGGCTG GAGCCCAATG CCCAGGCCCA GATGTACCGG 300
 CTGACCCTGC GCACCAGCAA GGAGCCCGTC TCCCGTCACC TGTGTGAGCT GCTGGCAGAN 360
 20 AGTTCTGAGC CTTGACTCT GCGCCGCGG ATGTGGCCGG CACTGGGCAG CCCCTTGGAC 420
 TGAGGCAGTT TTGCTGGATG GGGGACCTCC ACTGGTGACA GAGAAGACAC CAGGGTTTGG 480
 25 GGGATGCCTG GGACTTTCCT CCGGCTTTT GTATTTTAT TTTTGTTCAT CTGCTGCTGT 540
 TTACATTCTG GGGGGTTAGG GGGAGTCCCC CTCCCTCCCT TTCCCCCACA AGCACAGAGG 600
 GGAGAGGGGC CAGGGAAGTG GATGTCTCCT CCCCTCCAC CCCACCTGT TGTAGCCCTT 660
 30 CCTACCCCT CCCCATCCAG GGGTGTGTA TTATTGTGAG CGAATAACA GAGAGACGTT 720
 AACAGCCCCA TGTCTGTGTC CATCACCCAN TGNTAGGTAG TCAAAGAAGT GGGGTGAGGG 780
 35 CATGCAGAGT 790

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 776 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 117:

CAGCGCTGGA AGCAGCTGAG CCTGTGAGG GTGGGAGGG GCGGAGCAA AGCCGCGCCT 60
 CTGGGTGGGC GGGTCGGGCC GTCCAGGTCC CTGACTTGAA CCTTCCCGGT CCCAGCCCT 120
 CAACAGGAGG CGCAGAAAAT CTTCAAAGCC AACCACCCCA TGGACGCAGA AGTTACTAAG 180
 55 GCCAAGCTTC TGGGGTTTGG CTCTGCTCTC CTGGACAATG TGGACCCCAA CCTGAGAAC 240
 TTCGTGGGGG CGGGGATCAT CCAGACTAAA GCCCTGCAGG TGGGCTGTCT GTTTCGGCTG 300
 60 GAGCCCAATG CCCAGGCCCA GATGTACCGG CTGACCCTGC GCACCAGCAA GGAGCCCGTC 360

TCCCTCACC TGTGTAGCT GCTGTCACAG AGTTGTGAGC CCTGSACTCT GCCCCGGGG 420
 5 ATGTGGCCGG CACTGGGAG CCCCCTGAGC TGAGGCAGTT TTGGTGGATG GGGACCTCC 480
 ACTGCTGACA GAGAAGACAG CAGGCTTTGG GGGATGCCTG GGACTTTCTT CCGCCCTTTT 540
 GTATTTTAT TTTTGTTCAT CTGCTGCTGT TTACATTCTG GGGGTTAGG GGGAGTCCCC 600
 10 CTCCCTCCCT TCCCCCCCCA AGCAGAGAGG GGAGAGGGGC CAGGGAAGTG GATGCTCTCT 660
 CCCCTCCAC CCCACCTGT TGTAGCCCT CCTACCCCT CCCCATCCAG GGGCTGTGTA 720
 15 TTATTGTGAG CGAATAACA GAGAGACGCT TAAAAAATA AAAAAAAT TGAGGG 776

20 (2) INFORMATION FOR SEQ ID NO: 118:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 453 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 118:

30 GGTTCGACA CCAGATGTC TCTGCTCCTG GTTAATGTCA GTGAGGGCTG GAAGTTGAAT 60
 AAATGAGAAC AGGAGTGTC TGGGCCCATG TAAATGATCC TCCCTTGAAA GGAGGAACAG 120
 CTTTCATCAT TTGTTCCAGC TAAGCCTTGC ATGCATTATA GATCTGGTGC TAAGCAGTGG 180
 35 GAAAGATCTC ATAAGTAATG TTTATGTTT TTTCKGTCTC TCYTCTTCKG TTGTCTTGG 240
 CTGTGGGTT GTGTTTGKG TTGTTAACTG GAAAATTGCT ATAAGCCAGT TGTCYCKAAK 300
 40 TTTWAAAAAC GAATTAGAAA AACCATAAAA TCYTCTGGCC YATGCACATK GTCCCYGTTT 360
 TGTGAAAAACA TTAAAGGTA AATAAAAAGG AAGGAGAACA GTCAATAATG TGCATCAAAT 420
 ATATTCTGAG TTCTAGAGAA ATTAATGACC AAG 453

45

(2) INFORMATION FOR SEQ ID NO: 119:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2016 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

AGGCTGTTCA CAGGCACCCC GAGACAGCGT CCCCCCTCTG GCGCACTGG ATTTGACGTT 60
 60 GCAGGACGCG CGGCTGGAAC CCCAGGCCCG CGCTGCTCAC AGACCGGGAC TCCGCTCCG 120

	CTTCCCAGAG GCGTGGCGAG GCGCTGCGGG ANCCCAACAG GATGCGTTCC GTGCGTTCCA	180
	TCAAGATCTC AATTTTGTGC GCAATTCCTA CAGCCCCTGT TGATTGGAGA GCTGGCTCCG	240
5	GAAGAACCCA GCCAKGATGS ACCCCTGAAT GCGCATGGTC GAGGACTTCC GAGCCCCTGCA	300
	CCAGGCAGCC GAGGACATGA AGCTGTTTGA TGCCAGTCCC ACCTTCCTTG CTTTCCTACT	360
10	GGGCCACATC CTGGCCATGS AGGTGCTGGC CTGGCTCCTT ATCTACCTCC TGGGTCTTGG	420
	CTGGGTGCCC AGTGCCCTGS NCCGCTTCA TCCTGGCCAT CTCTCAGGCT CAGTCTTGGT	480
	GTCTGCAGCA TGACCTGGGC CATGCTCCAT CTTCAAGAAG TCCTGGTGA ACCACGTGGC	540
15	CCAGAAGTTC GTGATGGGGC AGCTAAAGGG CTTCTCCGCC CACTGGTGA ACTTCCGCCA	600
	CTTCCAGCAC CACGCCAAGC CCAACATCTT CCACAAAGAC CCAGACGTGA CGGTGGCGCC	660
20	CGTYTTCCTC CTGGGGGAGT CATCCGTCGA GTATGGNCAA GAAGAAACGC AGATACCTAC	720
	CCTACAACCA GCAGACCTG TACTTCTTCC TGATCGGCCC GCCGCTGCTC ACCCTGGTGA	780
	ACTTTGAAGT GGAAATCTG GCGTACATGC TGGTGTGCAT GCAGTGGGCG GATTGTCTCT	840
25	GGGCCGCCAG CTTCTATGCC CGCTTCTTCT TATCCTACCT CCCCTTCTAC GGCCTCCCTG	900
	GGGTGCTGCT CTTCTTTGTT GCTGTCAGGT ATGGCAGGGA GTGGCGAGGT CACACACAGG	960
30	CGACAGGTGA CCCCCACTGC AGCCCCCAC CAGAGCTTCC CTTTTCCTCT CTGCAGAATG	1020
	GGGCCAGTGG TACTGCCTCC CTGGCTTGCT GGTGGAATCA CATAAACACA AGYTTCAGGA	1080
	GCCCAGGGTC GGTGGGTTTA GGGAGCGTGG CCTGGCTTGT AAGTGGCCCC GTGGGTGTCTG	1140
35	GAGTGCTCTT GGA CTGAGCC TCACAGTGA CACTGCTCCA TTCAGATTCT TTAAACACTG	1200
	GCAAGGGGGC GATGGCCACA ATCCTATTGT ACAGATAAGG AAGTCAAGGC CAYTTGGGGA	1260
40	CAGYTGTCTT TCCAGCCTCC ACTCAGGGTG CCTTAAGTGG TGAGCTGGAC CTAGGGCAGT	1320
	GCCGAGCYTC CCCACAGGGT CCTGGAAGC CACTGGTTCG TGTGGATCAC ACAGATGAAC	1380
	CACATCCCCA AGGAGATCGG CCACGAGAAG CACCGGACT GGTCTAGCTC TCAGCTGGCA	1440
45	GCCACCTGCA ACGTGGAGCC CTCACCTTTC ACCAACTGGT TCAGCGGGCA CCTCAACTTC	1500
	CAGATCGAGC ACCACCTCTT CCCCAGGATG CCGAGACACA ACTACAGCCG GGTGGCCCCG	1560
50	CTGGTCAAGT CGCTGTGTGC CAAGCACGGC CTCAGCTACG AATGAAGCCC TTCCTCACCG	1620
	CGCTGGTGA CATCGTCAGG TCCTGAAGA AGTCTGGTGA CATCTGGCTG GACGCCTACC	1680
	TCCATCAGTG AAGGCAACAC CCAGCGGGC AGAGAAGGGC TCAGGGCACC AGCAACCAAG	1740
55	CCAGCCCCCG GCGGGATCGA TACCCCCAMC CCTCCACTGG CCAGCCTGGG GGTGCCCTGC	1800
	CTGCCCTCCT GGTACTGTG TCTTCCCCTC GGCCCCCTCA CATGTGTATT CAGCAGCCCT	1860
60	ATGGCCTTGG CTCTGGGCCT GATGGGACAG GGTAGAGGG AAGGTGAGCA TAGCACATTT	1920

TCCTAGAGCG AGAATGCGG GAAAGCTGT ATTTTATAT TAAAATACAT TCAGATGTAA 1980

AAAAAAAAA AAAAAANCT CGAGGGGGG CCCCCG 2016

5

(2) INFORMATION FOR SEQ ID NO: 120:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2136 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

15

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

20 GGGGACGGAG CCGCTGTCAA CTCTCCAACT CAGCTCAGCT GATCGGTTGC CGCCGCCGCC 60

GCCGCCAGAT TCTGGAGGCG AAGAACGCAA AGCTGAGAAC ATGGACGTTA ATATCGCCCC 120

ACTCCGCCCC TGGGACGATT TCTTCCCGGG TTCCGATCGC TTGCCCCGGC CGGACTTCAG 180

25 GGACATTTCC AAATGGAACA ACCGCGTAGT GAGCAACCTG CTCTATTACC AGACCAACTA 240

CCTGGTGGTG GCTGCCATGA TGATTTCCAT TGTGGGTTT CTGAGTCCCT TCAACATGAT 300

30 CCTGGGAGGA ATCGTGGTGG TGCTGGTGT CACAGGGTTT GTGTGGGCAG CCCACAATAA 360

AGACGTCCTT CGCCGGATGA AGAAGCGCTA CCCCACGACG TTCGTTATGG TGGTCATGTT 420

GGCGAGCTAT TTCCTTATCT CCATGTTTGG AGGAGTCATG GTCTTTGTGT TTGGCATTAC 480

35 TTTTCCTTTG CTGTTGATGT TTATCCATGC ATCGTTGAGA CTTCCGAACC TCAAGAACAA 540

ACTGGAGAAT AAAATGGAAG GAATAGGTTT GAAGAGGACA CCGATGGGCA TTGTCCTGGA 600

40 TGCCCTAGAA CAGCAGGAAG AAGGCATCAA CAGACTCACT GACTATATCA GCAAAGTGAA 660

GGAATAAACA TAACCTACCT GAGCTAGGGT TGCAGCAGAA ATTGAGTTGC AGCTTGCCCT 720

TGTCCAGACC TATKTTCTGC TTGCGTTTTT GAAACAGGAG GTGCACGTAC CACCCAATTA 780

45 TCTATGGCAG CATGCATGTA TAGGCCGAAC TACTATCAGC TCTGATGTTT CAGAGAGAAG 840

ACCTCAGAAA CCGAAAGAAA ACCACCACCC TCCTATTGTG TCTGAAGTTT CACGTGTGTT 900

50 TATGAAATCT AATGGGAAAT GGATCACAGC ATTTCTTTAA GGAATTTAAA AAAAATAAAA 960

GAATTACGGC TTTTACAGCA ACAATACGAT TATCTTATAG GAAAAAATAA ATCATTTGTA 1020

AGTATCAAGA CAATACGAGT AAATGAAAAG GGTGTTAAAG TAGATGACAT CATGTGTTAG 1080

55 CCTGTTCTTA ATCCCTAGA ATTGTAATGT GTGGATATA AATTAGTTTT TATTATTCTC 1140

TTAAAAATCA AAGATGATCT CTATCACTTT GCCACCTGTT TGATGTGCAG TGGAAACTGG 1200

60 TTAAGCCAGT TGTTCACTACT TCSYTTACAA ATATAAAGAT AGCTGTTTAG GATATTTTGT 1260

265

TACATTTTTG TAAATTTTTG AAATGCTAGT AATGTGTTTT CACCAGCAAG TATTTGTTGC 1320
 AAACCTAATG TCATTTTCCT TAAGATGGTT ACAGCTATGT AACCTGTATT ATTCTGGACG 1380
 5 GACTTATTA AATACAAACA GACAAAAAAT AAAACAAAAC TTGAGTTCTA TTTACCTTGC 1440
 ACATTTTTTG TTGTTACAGT GAAAAAATG GTCCAAGAAA ATGTTTGCCA TTTTTCATT 1500
 GTTCGTTTT TAACTGGAAC ATTIAGAAAG AAGGAAATGA ATGTGCATT TATTAATTCC 1560
 10 TTAGGGGCAC AAGSAGGACA ATAATAGCTG ATCTTTTGAA ATTGAAAAA CGTCTTTAGA 1620
 TGACCAAGCA AAAAGACTTT AAAAAATGGT AATGAAAATG GAATGCAGCT ACTGCAGCTA 1680
 15 ATAAAAAATT TTAGATAGCA ATTGTTACAA CCATATGCCT TTATAGCTAG ACATTAGAAT 1740
 TATGATAGCA TGAGTTTATA CATCTATTA TTTTCTCTCC CTTTCTCATG TTTTATAAA 1800
 TAGGTAATAA AAAATGTTTT GCCTGCCAAT TGAATGATT CGTAGCTGAA GTAGAAACAT 1860
 20 TTAGSTTTCT GTAGCATTAA ATTGTGAAGA CAACTGGAGT GSTACTTACT GAAGAACTC 1920
 TCTGTATCTC CTAGAATAAG AAGCAATGAT GTGCTGCTTC TGATTTTCT TGCATTTTAA 1980
 25 ATTCTCAGCC AACCTACAGC CATGATCTTT AGCAGAGTGA TATCACCATG ACTTCACAGA 2040
 CATGCTCTAG AATCTGTACC CTTACCCACA TATGAAGAAT AAAATTGATT AAAGTTAAA 2100
 AAAAAAWAA AAAAAMWAGG GGGGCCCGGT WCCAG 2136
 30

(2) INFORMATION FOR SEQ ID NO: 121:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

40

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

GCCCTAGTAT CTGGGCAGCT GTGCATGGAG ATAGCCAGAG GAAACATTTT TTTTCTTAAT 60
 45 GRATTTGGTGA CCACATTTTG TTGTTCTTGC CTCCTATTAT CCGTGCSCCTA TTTGCATSCCT 120
 GGTTCCTTCT ACAGTAGTTT ATGTAAATGT TGTTTTGTCC TTGTCGTTCT CAGTAGAATT 180
 50 GGTTCGTAA ACGAAACCTG GTCCTGTAAT TTCAGTATA 219

55

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1686 base pairs

(B) TYPE: nucleic acid

60

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

5	GCTGGAGATT CACATTTTAC CTGATTTCCT TCAATGCCCG CATGCCCGTC ATTGTGGATA	60
	AACCTCGTT CTATGACATG AAGAAAGTTT GGGAGGGATA TCCCATACAG AGTACTATCC	120
	CTCCCCAGTA TTGGTACTAC ATGATTGAAC TTTCCTTCTA CTGGTCCCTG CTCTTCAGCA	180
10	TTGCCTCTGA TGTCAAGCGA AAGGATTTC AAGAACAGAT CATCCACCAT GTGRCACCA	240
	TCAATCTCAT CAGCTTTTCC TGGTTTGCCA ATTACATCCG AGCTGGGACT CTAATCATGG	300
15	CTCTGCATGA CTCTTCCGAT TACCTGCTGG AGTCAGCCAA GATGTTTAAC TACCGGGAT	360
	GGAAGAACAC CTGCAACAAC ATCTTCATCG TCTTCGCCAT TGTTTTATC ATCACCCGAC	420
	TGGTCATCCT GCCCTTCTGG ATCCTGCATT GCACCCTGGT GTACCCACTG GAGTCTATC	480
20	CTGCCTTCTT TGGTATTAC TTCTTCAATT CCATGATGGG AGTTCTACAG CTGCTGCATA	540
	TCTTCTGGGC CTACCTCATT TTGCGCATGG CCCACAAGTT CATAACTGGG AAAGCTGGTA	600
25	GAAGATGAAC GCAWGCRCGG GNAAGAAACA GAGAGCTCAG AGGGGGAGGA GGCTGCAGCT	660
	GGGGGAGGAG CAAAGAGCCG GCCCTTAGCC AATGGCCACC CCATCCTCAA TAACAACCAT	720
	CGTAAGAATG ACTGAACCAT TATTCAGCT GCCTCCAGA TTAATGCATA AAGCCAAGGA	780
30	ACTACCCYGC TCCCTGCGCT ATAGGGTCAC TTAAAGCTCT GGGGAAAAG GAGAAAGTGA	840
	GAGGAGAGTT CTCTGCATCC TCCCTCCTTG CTTGTACCC AGTTGCCCTT AAACCAAATT	900
35	CTAACCAGCC TATCCCCAGG TAGGGGGACG TTGGTTATAT TCTGTTAGAG GGGGACGGTC	960
	GTATTTTCTT CCTACCCCG CAAGTCATCC TTCTACTGC TTTTGAGGCC CTCCCTCAGC	1020
	TCTCTGTGGG TAGGGGTTAC AATTCACATT CCTATTCTG AGAATTGGC CCCAGCTGTT	1080
40	TGCCCTTGAC TCCCTGACCT CCAGAGCCAG GGTGTGCCT TATTGTCCA TGTGTGGGCC	1140
	TCAATCTGCC AAAGCTGGAC CAAGGTTAAC CTTTCTAAGC TCCCTAAGTT GAGCCAGAAA	1200
45	CCAAAGCTGA GCTTTTAACT TTCTCCCTCT ATGACACAAA TGAATTGAGG GTAGGAGGAG	1260
	GGTGCACATA ACCCTTACCC TACCTCTGCC AAAAAGTGGG GGCTGTACTG GGGACTGCTC	1320
	GGATGATCTT TCTTAGTGCT ACTTCTTTCA GCTGTCCCTG TAGCGACAGG TTTAAGATCT	1380
50	GACTGCCTCC TCCCTTCTCT GGCCTTTTCC CCCTTCCCTC TTCTCTTCAG CTAGGTAGC	1440
	TGTTTGGAG TAGAATGGCA ACTAATCTA ATTTTATTT ATTAAATATT TGGGGTTTTG	1500
55	GTTTTAAAGC CAGAATTACG GTTAGCACCCT AGCATTTCAG CAGAGGGACC ATTTTAGACC	1560
	AAAATGTAAT GTTAATGGST TTTTTTTAA AATTAAAGA TTAATAAAA AATATTAAAT	1620
60	AAAACATGCC AATAAGTGTG AGACTATTAG GAATTGAGAA GGGGATCAA CTAATAAAC	1680

GAAGAG

1686

5

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 1211 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

15

CAGCCTGTGC CAGACGAGGA GGTGATTGAG CTGTATGGGG GTACCCAGCA CATCCCACTA 60

TACCAGATGA GTGGCTTCTA TGGCAAGGGT CCCTCCATTA AGCAGTTCAT GGACATCTTC 120

20

TCGCTACCGG AGATGGCTCT GTGTCTCTGT GTGGTGGACT ACTTCTTGGG CCACAGCCTG 180

GAGTTTGACC AAACATCTCT ACAAGGACGT GACGGACGCC ATCCGAGACG TGCATGTGAA 240

GGGCTCATG TACCAGTGGG TCGAGCAGGA CATGGAGAAG TACATCCTGA GAGGGGATGA 300

25

GACGTTTGCT GTCCTGAGCC GCCTGGTGGC CCATGGGAAA CAGCTGTTCC TCATCACCAA 360

CAGTCTTTTC AGCTTCGTAG ACAAGGGGAT GCGGCACATG GTGGGTCCCG ATTGGCGCCA 420

30

CTCTTCGATG TGGTCATTGT CCAGGCAGAC AAGCCAGCT TCTTCACTGA CCGGCGCAAC 480

TTTCAGAAA CTCTGATGAGA AGGGCTCACT TCAGTGGGAC CGGATCACCC GCTTGGAAAA 540

GGGCAAGATC TATCGGCAGG GAAACCTGTT TGACTTCTTA CGCTTGACGG AATGGCGTGG 600

35

CCCCCGCGTG CTCTACTTCG GGGACCACCT CTATAGTGAT CTGGCGGATC TCATGCTGCG 660

GCACGGCTGG CGCACAGGCG CCATCATCCC CGAGCTGGAG CGTGAGATCC GCATCATCAA 720

40

CACGAGCAG TACATGCACT CGCTGACGTG GCAGCAGGCG CTCACGGGGC TGCTGGAGCG 780

CATGCAGACC TATCAGGACG CGGAGTCGAG GCAGGTGCTG GCTGCCTGGA TGAAAGAGCG 840

GCAGGAGCTG AGGTGCATCA CCAAGGCCCT GTTCAATGCG CAGTTCGGCA GCATCTTCCG 900

45

CACCTTCCAC AACCCACCT ACTTCTCAAG GCGCCTCGTG CGCTTCTCTG ACCTCTACAT 960

GGCTCCCTC AGCTGCCTGC TCAACTACCG CGTGGACTTC ACCTTCTACC CACGCCGTAC 1020

50

GCCGCTGCAG CACGAGGCAC CCCTCTGGAT GGACCAGCTT CTGCACCGGC TGCATGAAGA 1080

CCCCCTTCTT TGGTGACATG GCCCACATCC GCTGAGGGCA CCTTTATTGT CTGGGACAGG 1140

CCCTCAGCCC CTCCTGCCCC ATCCACCCAG ACAAGCAATA AAAGTGGTCT CCTCCCTGAA 1200

55

AAAAAAAAA A 1211

60

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 1804 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 124:

10 CGCACCTATG GGCTCGCTAC CAGGACATGC GGAGACTGGT GCACGACCTC CTGCCCCCGG 60
 AGGTCTGCAG TCTCCTGAAC CCAGCAGCCA TCTACGCCAA CAACGAGATC AGCCTGCGTG 120
 15 ACGTTGAGGT CTACGGCTTT GACTACGACT ACACCCTGGC CCAGTATGCA GACGCACTGC 180
 ACCCCGAGAT CTTCACTACC GCCCCTGACA TCCTGATCGA GCACTACAAG TACCCAGAAG 240
 GGATTGGGAA GTATGACTAC AACCCAGCT TTGCCATCCG TGGCCTCCAC TATGACATTC 300
 20 AGAAGAGCCT TCTGATGAAG ATTGACGCCT TCCACTACGT GCAGCTGGGG ACAGCCTACA 360
 GGGGCCTCCA GCCTGTGCCA GACGAGGAGG TGATTGAGCT GTATGGGGGT ACCCAGCACA 420
 25 TCCCACTATA CCAGATGAGT GCCTTCTATG GCAAGGGTCC CTCCATTAAG CAGTTCATGG 480
 ACATCTTCTC GCTACCGGAG ATGGCTCTGC TGTCTGTGT GGTGGACTAC TTTCTGGGCC 540
 ACAGCCTGGN AGTTTGACCA AGCACATCTC TACAAGGACG TGACGGACGC CATCCGAGAC 600
 30 GTGCATGTGA AGGGCCTCAT GTACCACTGG ATCGAGCAGG ACATGGAGAA GTACATCCTG 660
 AGAGGGGATG AGACGTTTGC TGTCTGAGC CGCCTGGTGG CCCATGGGAA ACAGCTGTTC 720
 35 CTCATACCA ACAGTCCCTT CAGCTTCGTA GACAAGGGGA TGCGACACAT GGTGGGTCCC 780
 GATTGGCGCC ACTCTTGGAT GTGGTCATTG TCCAGGCAGA CAAGTCCAGC TTCTTCACTG 840
 ACCGGCGCAA GCTTTTCAGA AAACCTCGATG AGAAGGGCTC ACTTCAGTGG GACCGGATCA 900
 40 CCCGCTTGA AAAGGGCAAG ATCTATCGGC AGGGAACCT GTTTGACTTC TTACGCTTGA 960
 CGAATGGCG TGGCCCCCGC GTGCTCTACT TCGGGACCA CCTCTATAGT GATCTGGCGG 1020
 45 ATCTCATGCT GCGGCACGGC TGGCGCACAG GCGCATCAT CCCCAGCTG GAGCGTGAGA 1080
 TCCGCATCAT CAACACGAG CAGTACATGC ACTCGCTGAC GTGAGAGCAG GCGCTCACGG 1140
 GGCTGCTGGA GCGCATGAG ACCTATCAGG ACGCGGAGTC GAGGAGGTG CTGGCTGCCT 1200
 50 GSATGAAAGA GCGGACGAG CTGAGGTGCA TCACCAAGGC CCTTTCAAT GCGCAGTTTC 1260
 GCAGCATCTT CCGCACCTTC CACAACCCCA CCTACTTCTC AAAGTCGCCT CGTGCGCTTC 1320
 55 TCTGACCTCT ACATGGCCTC CCTCAGCTGC CTGCTCAACT ACCGCGTGA CTTCACCTTC 1380
 TACCCACGCC GTACGCCGCT GACGACGAG GCACCCCTCT GSATGGACCA GCTCTGCACC 1440
 60 GGCTGCATGA AGACCCCTT CCTTGGTGAC ATGGCCACA TCCGCTGAGG GCACCTTTAT 1500

TGTCTGGGAC AGGCCCTCAG CCCCTCCTGC CCCATCCACC CAGACAAGCA ATAAAAGTGG 1560
 TCTCCTCCCT GTGCATGCTT CTGCTTTCAG CCCCAGCCTC GTCACCTGAC TGTGAGGATC 1620
 5 CTCTGGGTGT CAGGGAAGTC CTCCTCCAGC AGTGAGTCAT CGAAGGGTTC ACAAAGGTG 1680
 TCGCTGCCAA AGACAGGGTT GGGACAGAG ACCAGGGTGG GGTGGTCCC TTCTTGCCAC 1740
 GGTGAGAAGT CGTCGTCAGC CGACGCGTG GGTGACCCG GGAATTCGG ACCGGTACCT 1800
 10 GCAG 1804

15

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 1282 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 125:

25 CCGCAGGNCA GCGACGCGAC TCTGGTGGG GCCGTCTTCT TCCCCCGAG CTGGGCGTGC 60
 GCGGCCGCAA TGAAGTGGGA GCTGCTGCTG TGGCTGCTGG TGCTGTGCGC GCTGCTCCTG 120
 30 CTCTTGGTGC AGCTGCTGCG CTTCCTGAGG GCTGACGGCG ACCTGACGCT ACTATGGGCC 180
 GAGTGGCAGG GACGACGCCC AGAATGGGAG CTGACTGATA TGGTGGTGTG GGTGACTGGA 240
 GCCTCGAGTG GAATTGGTGA GGAGCTGGCT TACCAGTTGT CTAAGTAGG AGTTTCTCTT 300
 35 GTGCTGTCAG CCAGAAGAGT GCATGAGCTG GAAAGGGTGA AAAGAAGATG CCTAGAGAAT 360
 GGCAATTAA AAGAAAAAGA TATACTTGTT TTGCCCTTG ACCTGACCGA CACTGGTTCC 420
 40 CATGAAGCGG CTACCAAAGC TGTCTCCAG GAGTTTGTA GAATCGACAT TCTGGTCAAC 480
 AATGGTGGAA TGTCCAGCG TTCTCTGTGC ATGGATACCA GCTTGGATGT CTACAGAAAG 540
 CTAATAGAGC TTAAGTACTT AGGGACGGTG TCCTTGACAA AATGTGTTCT GCCTCACATG 600
 45 ATCGAGAGGA AGCAAGGAAA GATTGTTACT GTGAATAGCA TCCTGGGTAT CATATCTGTA 660
 CCTCTTTCCA TTGGATACTG TGCTAGCAAG CATGCTCTCC GGGGTTTTTT TAATGGCCTT 720
 50 CGAACAGAAC TTGCCACATA CCCAGGTATA ATAGTTTCTA ACATTTGCCC AGGACCTGTG 780
 CAATCAAATA TTGTGGAGAA TTCCCTAGCT GGAGAAGTCA CAAAGACTAT AGGCAATAAT 840
 GGAGACCAGT CCCACAAGAT GACAACAGT CGTTGTGTGC GGCTGATGTT AATCAGCATG 900
 55 GCCAATGATT TGAAAGAAGT TTGGATCTCA GAACAACCTT TCTTGTTAGT AACATATTTG 960
 TGGCAATACA TGCCAACCTG GGCCTGGTGG ATAACCAACA AGATGGGGAA GAAAAGGATT 1020
 60 GAGAAGTTTA AGAGTGGTGT GGATGCAGAC TCTTCTTATT TTAAGTCTT TAAGACAAA 1080

CATGACTGAA AAGAGCACT CTACTTTTCA AGCCACTGGA GGGAAAAATG GAAAACATGA 1140
 AAACAGCAAT CTCTCTATGC TTCTGAATAA TCAAAGACTA ATTGTGRCY TTACTTTTTA 1200
 ATAGATATGA CTTTGCTTCC AACATGGAAT GAAATAAAAA ATAAATAATA AAAGATTGCC 1260
 ATGGAAAAAA AAAAGNNGG AN 1282

10

(2) INFORMATION FOR SEQ ID NO: 126:

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1296 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:

GGCAGAGCTT AGAGTGTGGA AAAGGCAACC AGGTTGGCCG TAAGTGCCTG CTGGAATGCG 60
 25 TGTGCCTCCA CAGGRTCTG GGCATCCGGA CTGATAACCA GCCGCCAGA CTGAGGGATG 120
 GAAGGCACTG AGATGGGGC CCGTCCAGGC GGACACCCGC AGAATGGAG CTTTCTGTGG 180
 TCTCTGCAC TCTGGCTGCC TCTTGCCCTC TCTGTGTCTC TCTTCTTGG TCTCTCCCTC 240
 30 TCTCTCCTC AGCCTGGTCT TTCTCTTGG TGCACACTTA GTTATTGTG TGAGCAATGG 300
 AAGTTCAAAG GAACTCCCTC TCCAGCTCTT CTGAATCTTG GACACAGCC TAAAAAGGAC 360
 35 AAAAAAGTAG AAGACAGCAT AGCAACTCAG CTCAGGGRGC TACCAGAGAA AAATAGCAAC 420
 TGATGTGGGT GCTTTTTTTT TTTTTTAAAT TTGAATAAAA AGAATTAGAA GTGATGTCTT 480
 TTTATAAAAT GCTTCTCTCC CCTTCCCGCC TACAGTCTCT TCTCTCTCCC TTAGAGGGGG 540
 40 GAAAGTGTAT AAACCTACAG GGTGTGTAGT CTGAAAAGAG GATCCCCCTC ACCCCCACCC 600
 TGGGCAGAGC AGTGGGGTT GGGGGTGGG AGAGGGGAC ACAGATCCTG GCACACTGTG 660
 45 GATATTTCTT GAGAGTTGCA GTCTCTTGTG GCCCAAACAG GTTAGGTAGA CTATCGCCTT 720
 TGGCAGGTGC CACCTTTTGG TACCAACATG TTCTGAGGTG TTAGGATTG GGTGGGTTT 780
 TTTTGTGTTG TTTTTTTTTT CCNNTTGGTC TTTTTTTTTT TTYCCTKTA AAGAAAAGCT 840
 50 AAAGGCCGCT GTAGTCTCTG GTGGCAGGCT CTCCATGGAT GTAGCATATC GAAGATAATT 900
 TTTTACTGTC ATTTTATGG ATTATTTTGT AATGTGTGAT TCCGTCTGCT GAGGAGGTG 960
 55 GAGGGGCTCC AGGAAAGCC ACCACCTTC AGTGAGGTTG CTCCCCAGCT GAGCGCACCG 1020
 GGCATGGGAT GTGAGGGTG GCGACACACC CTGTGCCTCT CCAAGGCTGG GCGCGTGGG 1080
 CGTCCAGAGT CTCTCTGGT CTCAGATGTC CATCTGCCAC CTCTGTATA GGCTCTAGCT 1140
 60

271

AGAAGGGAGG CTGAGGCTAG AAGAAAGTTA TTCCCGAAGA AAAAAAGAAT GAAAAGTCAT 1200
 TGTACTGAAC TGTTTTATA TTTTAAAAG TTAATTTTA AAGCGGACGT CGTGGGTGGA 1260
 5 CCCGGGAATT CCCG3ACCGG TACTGTCAGG TCTAAC 1296

10 (2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 737 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:

20 GGCANAGTGG AGGCAATGCC AGCTCCAGGA CAGAGGCTCA GGTGCCCAAC GGGCAAGGCA 60
 GCCCAGGGGG CTGTGTCTGT TCAAGTCAGG CTTCCCCGGC CCYTCGCGCA NCAGCGCTTC 120
 CACGGGCAGC CCGGGGCCCC ACCCCACGCA CTGAAGAGGC CGCCTGGGCT GCCATGGCCC 180
 25 TGACCTTCCT GCTGGTGTG CTCACCCTGG CCACGCTCTG CACACGGCTG CACAGAACT 240
 TCCGACGCGG GAGAGCATC TACTGGGGG CCACAGCGGA CAGCCAGGAC ACAGTGGCTG 300
 30 CTGTGCTGAA GCGGAGGCTG CTGCAGCCCT CGC3CCGGGT CAAGCGCTCG CGCCGGAGAG 360
 CCYTCYTCCT GCCCAGCGCG GACAGCGGCC CGGAAGGCGA GAGCTCGGAG TGACGGCCTG 420
 GGACCTGCCA CTGTGGCGTG CCGTCTCCCC GCGCCGCGAG GCCCGGAMCT NTGCCACGTG 480
 35 GACCGCGCGC NGGCGCTMC CCTGGTGGCG ATGCGCGGC ACTGGCGAGC ACTGCGKGGC 540
 CTTTCTCTCT TGTGGTTGC TGAGTGGGCG GCCAAGGGGA GAAAGGAGC CGCTTYTGCC 600
 40 TCCCTTGCCA AAAGTCCGTT TCTAATTAAA TTATTTT TAGAAAAAAA AAAAAAAAAA 660
 AAAAAAAAAA AAAAAAAAAA AAAAAAAC TCGAGGGGGG GCCCGGTACC CAATTNGCCA 720
 AATAGCGATC GTATNAA 737
 45

50 (2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1925 base pairs
 (B) TYPE: nucleic acid
 55 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 128:

60 CCCCCCTCC AAAGCTAACC CTCGGGCTTG AGGGGAAGAR GCTGACTGTA CGTTCTTCT 60

	ACTCTG99AC CACTCTCCAG GCTGCCATGG G9CCCAGCAC CCTCTCCTC ATCTTGTTC	120
	TTTTGTGATG GTGGGAGCC CTCCAAGGAC AGGAGCACCA CTTTGTGGAG TACATGGAAC	180
5	GCGGACTAGG TGCTTTAGAG GAACGGGTGG CCGAGTGCCA G3ACCAGAGT AGTCGGCATG	240
	CTGCTGAGGT GGGGACTTC AAGAACAAGA TGCTGCCACT GTTG3AGGTG GCAGAGAAGG	300
10	AGCGG3AGGC ACTCAGAACT GAGGCCGACA CCACTCTCCG GAGAGTGGAT CGTCTGGAGT	360
	GGGAGGTAGA CTATCTGGAG ACCCAGAACC CAGCTCTGCC CTGTGTAGAG TTTGATGAGA	420
	AGGTGACTG3 AGGCCCTGGG ACCAAAGGCA AGGGAAGAAG GAATGAGAAG TACGATATGG	480
15	TGACAGACTG TGGCTACACA ATCTCTCAAG TGAGATCAAT GAAGATTCTG AAGCGATTTG	540
	GTGGCCGAGC TGGTCTATGG ACCAAGGATC CACTGGGGCA AACAGAGAAG ATCTACGTGT	600
20	TAGATGG3AC ACAGAATGAC ACAGCCTTTG TTTTCCAAG GCTGCTGAC TTCACCTTTG	660
	CCATG3CTGC CGGAAAGCT TCCCGAGTCC G3GTGCCCTT CCCCCTGGGTG GGCACAGGGC	720
	AGCTG3TATA TGSTGCTTT CTTIATTTTG CTGG3AGGCC TCTTGAAGA CCTGCTGGAG	780
25	GTGGTGAGAT GGAGAACT TTGCAGCTAA TCAATTTCCA CTTG3CAAAC CGAACAGTGG	840
	TGGACAGCTC AGTATTCCCA GCAGAGGGGC TGATCCCCC CTACGGCTTG ACAGCAGACA	900
30	CCTACATCGA CCTGGCAGCT GATGAG3AAG GTCTTTGGGC TGTCTATGCC ACCCGGGAGG	960
	ATGACAG3CA CTGTGTCTG GCCAAGTTAG ATCCACAGAC ACTG3ACACA GAGCAGCAGT	1020
	GGGACACACC ATGTCCAGA GAGAATGCTG AG3CTGCCTT TKTCATCTGT G3GACCTCT	1080
35	ATGTGCTCTA TAACACCCGT CCTGCCAGTC G3CCCCGAT CCAGTGCTCC TTTGATGCCA	1140
	GCGGACCTG ACCCCTGAAC GGGCAGCACT CCTTIATTTT CCCC3CAGAT ATG3TGCCCA	1200
40	TGCCAGCTC CGCTATAACC CCCGAGAAG CCA3CTCTAT GCCTGGGATG ATGGCTACCA	1260
	GATTGTCTAT AAGCTGGAGA TGAGGAAGAA AGAGGAGGAG GTTTGAGGAG CTAGCCTTGT	1320
	TTTTTGATC TTCTCACTC CCATACATT ATATATATC CCCACTAAAT TTCTTGTTC	1380
45	TCATTCTTCA AATGTGGGCC AGTTGTGGCT CAAATCCTCT ATATTTTGTAG CCAATGGCAA	1440
	TCAAATCTT TCAGCTCCTT TGTTCATAC G3AACTCCAG ATCCTGAGTA ATCCTTTTAG	1500
50	AGCCCGAAGA GTCAAAACCC TCAATGTTCC CTCTGTCTT CCTGCCCAT GTCAACAAAT	1560
	TTCA3GCTAA G3ATGCCCA GACCCAGGGC TCTAACCTTG TATGCGGGCA GGCC3AGGGA	1620
	GCAGGCAGCA GTGTTCTTCC CCTCAGAGTG ACTTGG3GAG GGAGAAATAG GAGGAGAGT	1680
55	CCAGCTCTGT CCTCTCTTCC TCACTCCTCC CTTCAGTGTG CTGAGGAACA GGACTTTCTC	1740
	CACATTGTTT TGTATTGCAA CATTTTGCAT TAAAAGGAAA ATCCAMAAAA AAAAAAAAAA	1800
60	AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1860

ACTGGGKCG CTGTCCCTTC TGTCGTCTTC TGCAGTCGT ACCCTTCTGT CGTCTTCTCG 1920
CAGCC 1925

5

(2) INFORMATION FOR SEQ ID NO: 129:

10

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 129:

TCCTAGCTTC CCAACCTCT GGCATCCCCA GCACTGATGG TCCTGGCATC CACGGCTGAG 60
GCCAGCCGTG ACTGCTTCCA TCCCTTGTC A GACCCACGA CCCTTGGTG TACCTGTYTC 120
AGTTGACAAG GACGTGCATA TTCCTTTCAC CAACGGTTCC TATACCTTTG CCTCTATGTA 180
CCATCGGCAA GGTGGGGTGC CAGGCACTTT TGCCAATCGT GATTTCCTCC CTCTCTACT 240
ACACCTCCAC CCTCAATTTG CTCCCCCAA TCTAGATTGC ACCCAATCA GTATGCTGAA 300
TCATAAGTGG TGTGGGGTTC TCCGGCCTTT GCTCCACCC GGGGACCGGG RGAGYTATCA 360
GTCAGTTTTC CGCCGGCCAA GCGACTTAAG AACTGTCATG ACACAGAGTC TCCCCACTTG 420
CGCTCTCAG ATGCAGATGG GAANGAATAT GACTTTGGGA CACAGCTGCM ATCTAGCTCC 480
CCCGGTTTTC TAAAGGTTGA TGACACTGGG AAGAAGATTT TTGCTGTCTC TGGCCTCAT 540
TCTGATCGGG AAGCCTCATC TAGCCAGAG GNTCGGNAAT GACAGATGTA AGAAGAAAGC 600
AGCGGCATTG TTCGACAGCC AGGCCCCAAT TTGCCCCATC TGCCAGGTCC TGCTGAGGCC 660
CAGTGAGCTG CAGGAGCATA TGGAGCAGGA ACTGGAGCAG CTAGCCCAAC TGCCCTCGAG 720
CAAGAATTCC CTCTGAAGG ATGCCATGGC TCCAGGCACC CCAAAGTCCC TCCTGTTGTC 780
TGCTTCCATC AAGAGGGAAG GAGAGTCTCC AACGGCATCA CCCCCTCAT CTGCCACCGA 840
TGACCTCCAC CATTCAGACA GATACCAGAC CTTTCTGCGA GTACGAGCCA ACCGGCAGAC 900
CCGATTGAAT GYTCGGATTG GGAAAATGAA ACGGAGGAAG CAAGATGAAG GGCAGGTATG 960
TCCCTGTGTC AACCGCCCCC TGGCAGGATC GAGCAGGAG ATGAGTAGGC ATGTGGAGCA 1020
TTGCCTTTCT AAGAGGGAAG GCTCCTGCAT GGCTGAGGAT GATGCTGTGG ACATCGAGCA 1080
TGAGAACAAC AACCGCTTTG AGGAGTATGA GTGTGTGGA CAGAAGCGGA TACGGGCCAC 1140
CACTCTCTG GAAGGTGGCT TCCGAGGCTC TGCTTTCATC ATGTGCAGCG GCAAAGAGAA 1200
CCCGACAGT GATGCTGACT TGGATGTGGA TGGGATGAC ACTCTGGAGT ATGGGAAGCC 1260
ACAATACACA GAGGCTGATG TCATCCCCTG CACAGGCGAG GAGCCTGGTG AAGCCAAGGA 1320

	GAGAGAGGCA CTTCGGGGCG CAGTCCCTAA TGCGGGCCCT CCCAGCACGC GCATCACACC	1380
5	TGAGTTCTCT AAATGGGCCA GTGATGAGAT GCATCCACC AGCAATGCTG AAAGCAGCAA	1440
	GCAGGAGGCC ATGCAGAAGA CCTGCAAGAA CAGCGACATC GAGAAAATCA CCGAAGATTC	1500
	AGCTGTGACC ACSTTTGAGG CTCTGAAGGC TCGGTCAGA GAACCTGAAC GGCAGCTATC	1560
10	TCTGCGGGAC CGTTACAAAT GCCTCATCTG CATGGACTCG TACTCGATGC CCTAACGTC	1620
	CATCCAGTGT TGGCAGTGC ACTGCGAGGA GTGCTGGCTG CGGACCCCTG GTGCCAAGAA	1680
15	GCTCTGCCCT CAGTGCAACA CGATCACAGC GCCCGAGAC CTGCGGAGGA TCTACTTGTG	1740
	AGCTATCTGC CCCAGGCAGG CCTCGCCTCC AGCAGCCCCA CCTGCCCCCA GCTCTGTGA	1800
	CAGTGACCGT YTCCTTTTGT ACATACTTGC ACACAGGTTC CCCATGTACA TACATGCACA	1860
20	TACTCAAACA TGCGTACACA CACACACATT TACACACGCA GGACTCTGGA GTCAGAGTAG	1920
	AGGCTGTGGC CCAGGCACTA CCTGCTGGCT CCCACCTATG GTTTGGGGGC CATACTGTTC	1980
25	CCAGCTCTGT TCCCAGGGTG GGGCAGGGAG GTGGGGTTG GGGGAGTAGT GGGGCACGGC	2040
	TCCTAAGATC CAGCCCCCAT ACTGACAGAC GGACAGACAG ACATGCAAAC ACCAGACTGA	2100
	AGTACATGTA ATATAGACCG TGTATGTTTA CAATGTTGTG TATAAATGGG ACAACTCCTC	2160
30	GCCCTCTACC TGTCCCTCC CCTTTGGTT GTATGATTTT CTCTTTTTC AAGAACCCT	2220
	GGAAGCAGCG CCTCCTTCAG GGTGCGCTGG GAGCTCGGCC CATCCACCTC TTGGGGTAYC	2280
35	TGCTCTCTC TCTCTGTGG TGTCCTTCC CTCTCCCATG TGCTCGGTGT TCAGTGGTGT	2340
	ATATTTCTTC TCCCAGACAT GGGCAGACG CCGCAAGGA CATGATCCTC TCCTTAGTCT	2400
	TAGCTCATGG GGCTCTTAT AAGGATTTGG GGGGTAGAGG CAGGAAATGG GAACCGAGCT	2460
40	GAAGCAGAGG CTGAGTTAGG GGGCTAGAGG ACAGTGCTCC TGGCCACCCA GCCTCTGCTG	2520
	AGAACCATTC CTGGGATTAG AGCTGCCTTT CCCAGGGAAA AAGTGTGTC TCCCCGACCC	2580
45	TCCCGIGGGC CCTGTGGTGT GATGCTGTGT CTGTATATTC TATACAAAGG TACTTGTCTT	2640
	TTCCCTTGT AACTACATT TGACATGGAT TAAACCAGTA TAAACAGTTA AAAAAAAAAA	2700
	AAAAAAAACT CGA	2712

50

(2) INFORMATION FOR SEQ ID NO: 130:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1011 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 130:

	AGAGGACGGT GTGACCCGGG AGGAAGTAGA GCGTGAGGAG GGTGAAGAAG GCATCTCTGA	60
5	GCAACCCTGC CCAGCTGACA CAGAGGTGGT GGAAGACTCC TTGAGGCAGC GTAAAAGTCA	120
	GCATGCTGAC AAGGGACTGT AGATTTAATG ATGCGTTTTT AAGAATACAC ACCAAAACAA	180
10	TATGTCAGCT TCCCTTTGGC CTGCAGTTTG TACCAAATCC TTAATTTTTY YTGAAATGAGC	240
	AAGCTTCTCT TAAAAGATGC TCTCTAGTCA TTTGGTCTCA TGGCAGTAAG CCTCATGTAT	300
	ACTAAGGAGA GTCTTCCAGG TGTGACAATC AGGATATAGA AAAACAAACG TAGTGTNTGG	360
15	GATCTGTTTG GAGACTGGGA TGGGAACAAG TTCATTTACT TAGGGGTCAG AAGTCTCGA	420
	CCAGAGGAGG CCATTCCCAG TCCTAATCAG CACCTTCCAG AGACAAGGCT GCAGGCCCTG	480
20	TGAAATGAAA GCCAAGCAGG AGCCTTGGCT CTGAGNCATC CCCAAAGTGT AACGTAGAAG	540
	CCTTGCATCC TTTTCTTGTG TAAAGTATTT ATTTTGTCA AATTGCAGGA AACATCAGGC	600
	ACCACAGTGC ATGAAAAATC TTTCACAGCT AGAAATTGAA AGGGCCTTGG GTATAGAGAG	660
25	CAGCTCAGAA GTCATCCCAG CCCTCTGAAT CTCTGTGCT ATGTTTATT TCTTACCTTT	720
	AATTTTCCA GCATTTCCAC CATGGGCATT CAGGCTCTCC AACTCTTCA CTATTATCTC	780
30	TTGCTCAGAG GACTCCAATA ACAGCCAGGT TTACATGAAC TGTGTTTGTT CATCTGACC	840
	TAAGGGGTTT AGATAATCAG TAACCATAAC CCCTGAAGCT GTGACTGCCA AACATCTCAA	900
	ATGAAATGTT GTRGCCATCA GAGACTCAAA AGGAAGTAAG GATTTTACAA GACAGATTAA	960
35	AAAAAATTG TTTTGTCCAA AAAANAAAAA AAAAAAATC GAAGGGGGGG C	1011

40 (2) INFORMATION FOR SEQ ID NO: 131:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2278 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 131:

50	GTAATTCGGC ACGAGGCGCC CAACATGGCG GGTGGGCGCT GCGGCCCGCA SCTAACGGCG	60
	CTCTTGGCCG CCTGGATCGC GGCTGTGGCG GCGACGGCAG GCCCCGAGGA GGCCGCGCTG	120
	CCGCCGAGC AGAGCCGGGT CCAGCCCATG ACCGCCTCCA ACTGGACGCT GGTGATGGAG	180
55	GGCGAGTGA TGCTGAAATT TTAGSCCCCA TGGTGTCCAT CCTGCCAGCA GACTGATTCA	240
	GAATGGGAGG CTTTGTGAAA GAATGGTGAA ATACTTCAGA TCAGTGTGGG GAAGGTAGAT	300
60	GTCATTCAAG AACCAGGTTT GAGTGGCCGC TTCTTTGTCA CCACTCTCCC AGCATTTTTT	360

	CATGCAAAGS ATGGGATATT CCGCGGTIAT CGTGCGCCAG GAATCTTCGA AGACCTGCAG	420
5	AATATATCT TAGAGAAGAA ATGGGAATCA CTCGAGCGTC TGACTGCGTG GAAATCCCCG	480
	GCTTCTCIAA CGATCTCTGS AATGCGTGGT CTTTMTAGCA TCTCTGCGAA GATATGSCAT	540
	CTTCACAAC ATTTACAGT GACTCTTGA ATTCTCTGT GCTCTCTTA TCTCTTTTC	600
10	CTCATAGCCA CTTGCTTTT TGGCTTTT ATGGCTCTGS TCTGCTGGT AATATCAGAA	660
	TCTTCTATG TGCACTTCC AAGGCATTTA TCTGAGCGT CTGAGCAGAA TCGGAGATCA	720
15	GAGSAGGCTC ATAGAGCTGA ACAGTTGCG GATCGCGAGS AGSAAAAAGA TGATTCAAAT	780
	GAAGAAGAAA ACAAGACAG CTTCTAGAT GATGAAGAAG ASAAAGAAGA TCTTGGCGAT	840
	GAGGATGAAG CAGAGGAAGA AGAGGAGSAG GACAACCTGS CTGCTGCTGT GSATGAGGAG	900
20	AGAACTGAGS CCAATGATCA GCGCGCCCA GAGAGGAGS GTCTGACCCG GGAGGNAAGT	960
	AGAGCCTGAG GAGGTTGAAG AAGGCATCTC TGAGCAACCC TGCGCAGCTG ACACAGAGGT	1020
25	GCTGGAAGAC TCCTTGAGGC AGCTTAAAG TCAGCATGCT GNCAGGGAC TGTAGATTTA	1080
	ATGATGCGTT TTCAAGAATA CACACCAAAA CAATATGTCA GCTTCCCTTT GGCTGCAGT	1140
	TCTACCAA TCTTAATTT TCTCTGAATG AGCAAGCTTC TCTTAAAGA TCTCTCTAG	1200
30	TCATTTGGTC TCATGCGAGT AAGCTCATG TATACTAAGG AGACTCTCC AGGTGTGACA	1260
	ATCAGGATAT AGAAAAACAA ACCTAGTGTN TGGGATCTGT TTGAGACTG GGATGGGAAC	1320
35	AAGTTCATTT ACTTAGGGT CAGAGAGTCT CGACCAGAGS AGGCCATTC CAGTCTAAT	1380
	CAGCACCTTC CAGAGACAAG GCTGCGAGGC TGTGAAATGA AAGCCAAGCA GGAGCCTTG	1440
	CTCTGAGCA TCCCCAAGT GTAAGTAGA AGCCTTGCAT CTTTCTCTG TGTAAAGTAT	1500
40	TTATTTTGT CAAATTGCAG GAAACATCAG GCACCACAGT GCATGAAAAA TCTTTCACAG	1560
	CTAGAAATG AAAGGCGCTT GGTATAGAG AGCAGCTCAG AAGTCATCCC AGCCTCTGA	1620
45	ATCTCTGTG CTATGTTTA TTTCTTACCT TTAATTTTC CAGCATTTCC ACCATGGGCA	1680
	TTGAGGCTCT CCACACTCT CACTATTATC TCTTGTGAG AGGACTCCAA TAACAGCCAG	1740
	GTTTACATGA ACTGTCTTG TCTATCTGA CTTAAGGCT TTAGATAATC ATAAACATA	1800
50	ACCGCTGAAG CTCTGACTG CAAACATCTC AAATGAAATG TGTGCGCAT CAGAGACTCA	1860
	AAAGGAAGTA AGSATTTTAC AAGACAGATT AAAAAAAT TCTTTTCTCC NAAATATAG	1920
55	TTTGTGTGA TTTTCTTA AGTTTCTAA GCAATATTT TCAAGCCAGA AGTCTCTAA	1980
	GTCTTGCAG TACAAGGTAG TCTGTGAAG AAAAGTTGAA TACTGTTTTG TTTTCATCTC	2040
	AAGGGTTCC CTGGTCTTG AACTACTTTA ATAATACTA AAAAACCCT TCTGATTTTC	2100
60	CTTCAGTGAT GTGCTTTTG TGAAGAATT AATGAACCTC ATACCTGAA AGTGAAAGAT	2160

TTGATTTTGT TTCCATCTTC TGTAATCTTC CAAAGAATTA TATCTTTGTA AATCTCTCAA 2220
 TACTCAATCT ACTSIAAGTA CCCAGGGGCG STAATTTCTT TAAAAAAAAA AAAAAAAA 2278
 5

10 (2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1088 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 132:
 GGCAGGGGCG GCGTGAACCC GTCGGGCACT GTGTCCCTGA CAA TGGAAC AGCCGACAGT 60
 20 GATGAGATGG CCCCAGAGCC CCACAGCACA CCCACATCGA TGTGCACATC CACCAGGAGT 120
 CTGCCCTGGC CAAGCTCCTG CTCACCTGCT GCTCTGCGCT GCGGCCCGG GCCACCCAGG 180
 25 CCAGGGGCAG CACCCGGCTG CTGGTGGCCT CGTGGGTGAT GCAGATCGTG CTGGGGATCT 240
 TGAGTGCAGT CCTAGGAGGA TTTTCTACA TCCGCGACTA CACCCTCCTC GTCACCTCGG 300
 GAGCTGCCAT CTGACAGGG GCTGTGGCTG TGCTGGCTGG AGCTGCTGCC TTCATTTAYG 360
 30 AGAAACGGGG TGCTACATAC TGGGCCCTGC TGAGGACTCT GCTARCGCTG GCAGCTTTCT 420
 CCACAGCCAT CCGTGCCTC AAACCTTTGA ATGAAGATTT CCGATATGGC TACTCTTAT 480
 35 ACAACAGTGC CTGCCGCATC TCCAGCTCGA GTGACTGGAA CACTCCAGCC CCCACTCAGA 540
 GTCCAGAAGA AGTCAGAAGG CTACACCTAT GTACCTCCTT CATGGACATG CTGAAGGCCT 600
 TGTTCAGAAC CCTTCAGGCC ATGCTCTTGG GTGTCTGGAT TCTGCTGCTT CTGGCATCTC 660
 40 TGGCCCCCTC GTGGCTGTAC TGCTGGAGAA TGTTCCTAAC CAAAGGGAAA AGAGACCAGA 720
 AGGAAATGTT GGAAGTGAGT GGAATCTAGC CATGCTCTC CTGATTATTA GTGCCTGGTG 780
 45 CTTCTGCACC GGGCGTCCCT GCATCTGACT GCTGGAAGAA GAACCAGACT GAGGAAAAGA 840
 GGCTCTTCAA CAGCCCCAGT TATCCTGGCC CCATGACCGT GGCCACAGCC CTGCTCCAGC 900
 AGCACTTGCC CATTCTTAC ACCCCTTCCC CATCTGCTC CGCTTCATGT CCCCTCCTGA 960
 50 GTAGTCATGT GATAATAAAC TCTCATGTTA TTGTTCCNAA AAAAAAAAAA AAAAAAAAT 1020
 TGGGGGGGGG CCGGTACCCA TTGGGCCTNN GGGGGNGGTT TAAATTAAT GGGGGGGGTT 1080
 55 TAAAAGGG 1088

60 (2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 553 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 133:

10 GGCAGAGAGC AGATGGCCTT GACACCAGCA GGGTGACATC CGCTATTGCT ACTTCTCTG 60
TCCCCACAG TTCTCTGGA CTCTCTGGA CCACAGTCCT CTGCCAGACC CCTGCCAGAC 120
CCCACTCCAC CATGATCCAT CTGGGTGACA TCCTCTTCTT GCTTTTGCTC CCAGTGGCTC 180
15 CAGCTCAGAC GACTCCAGGA GAGAGATCAT CACTCCCTGC CTTTACCCT GGCATTTCAG 240
GCTCTTGTC CGGATGTGGG TCCCTCTCTC TGCCGCTCCT GGCAGGCCTC GTGGGTGCTG 300
20 ATGCGGTGGC ATCGCTGCTC ATCGTGGGG CGGTCTTCTT GTGGGCACGC CCACGCCGCA 360
GCCCCGCCCA AGATGGCAAA GTCTACATCA ACATCCAGG CAGGGGCTGA CCTCTCTGCA 420
GCTTGACCT TTGACTTCTG ACCCTCTCAT CCTGATGGT GTGTGGTGGC ACAGGAACCC 480
25 CCGCCCCAAC TTTGGATTG TAATAAACA ATTGAAACAC CAAAAAAAAA AAAAAAAAAA 540
AAAAAAAAA AAA 553

30

(2) INFORMATION FOR SEQ ID NO: 134:

35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 467 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 134:

40

Met Arg Pro Gln Glu Leu Pro Arg Leu Ala Phe Pro Leu Leu Leu Leu
1 5 10 15

45

Leu Leu Leu Leu Leu Pro Pro Pro Cys Pro Ala His Ser Ala Thr
20 25 30

Arg Phe Asp Pro Thr Trp Glu Ser Leu Asp Ala Arg Gln Leu Pro Ala
35 40 45

50

Trp Phe Asp Gln Ala Lys Phe Gly Ile Phe Ile His Trp Gly Val Phe
50 55 60

Ser Val Pro Ser Phe Gly Ser Glu Trp Phe Trp Trp Tyr Trp Gln Lys
65 70 75 80

55

Glu Lys Ile Pro Lys Tyr Val Glu Phe Met Lys Asp Asn Tyr Pro Pro
85 90 95

60

Xaa Phe Lys Tyr Glu Asp Phe Gly Pro Leu Phe Thr Ala Lys Phe Phe
100 105 110

	Asn	Ala	Asn	Gln	Trp	Ala	Xaa	Ile	Phe	Gln	Ala	Ser	Gly	Ala	Lys	Tyr	
			115					120					125				
5	Ile	Val	Leu	Thr	Ser	Lys	His	His	Glu	Gly	Phe	Thr	Leu	Trp	Gly	Ser	
		130					135					140					
	Glu	Tyr	Ser	Trp	Asn	Trp	Asn	Ala	Ile	Asp	Glu	Gly	Pro	Lys	Arg	Asp	
	145					150					155				160		
10	Ile	Val	Lys	Glu	Leu	Glu	Val	Ala	Ile	Arg	Asn	Arg	Thr	Asp	Leu	Arg	
				165						170					175		
	Phe	Gly	Leu	Tyr	Tyr	Ser	Leu	Phe	Glu	Trp	Phe	His	Pro	Leu	Phe	Leu	
15				180					185				190				
	Glu	Asp	Glu	Ser	Ser	Ser	Phe	His	Lys	Arg	Gln	Phe	Pro	Val	Ser	Lys	
		195						200					205				
20	Thr	Leu	Pro	Glu	Leu	Tyr	Glu	Leu	Val	Asn	Asn	Tyr	Gln	Pro	Glu	Val	
	210						215					220					
	Leu	Trp	Ser	Asp	Gly	Asp	Gly	Gly	Ala	Pro	Asp	Gln	Tyr	Trp	Asn	Xaa	
	225				230						235				240		
25	Thr	Gly	Phe	Leu	Ala	Trp	Leu	Tyr	Asn	Glu	Ser	Pro	Val	Arg	Gly	Thr	
				245					250						255		
	Val	Val	Thr	Asn	Asp	Arg	Trp	Gly	Ala	Gly	Ser	Ile	Cys	Lys	His	Gly	
30				260				265					270				
	Gly	Phe	Tyr	Thr	Cys	Ser	Asp	Arg	Tyr	Asn	Pro	Gly	His	Leu	Leu	Pro	
		275						280					285				
35	His	Lys	Trp	Glu	Asn	Cys	Met	Thr	Ile	Asp	Lys	Leu	Ser	Trp	Gly	Tyr	
		290					295					300					
	Arg	Arg	Glu	Ala	Gly	Ile	Ser	Asp	Tyr	Leu	Thr	Ile	Glu	Glu	Leu	Val	
	305				310					315					320		
40	Lys	Gln	Leu	Val	Glu	Thr	Val	Ser	Cys	Gly	Gly	Asn	Leu	Leu	Met	Asn	
				325						330					335		
	Ile	Gly	Pro	Thr	Leu	Asp	Gly	Thr	Ile	Ser	Val	Val	Phe	Glu	Glu	Arg	
45				340				345					350				
	Leu	Arg	Gln	Met	Gly	Ser	Trp	Leu	Lys	Val	Asn	Gly	Glu	Ala	Ile	Tyr	
		355						360					365				
50	Glu	Thr	His	Thr	Trp	Arg	Ser	Gln	Asn	Asp	Thr	Val	Thr	Pro	Asp	Val	
		370					375					380					
	Trp	Tyr	Thr	Ser	Lys	Pro	Lys	Glu	Lys	Leu	Val	Tyr	Ala	Ile	Phe	Leu	
	385				390					395					400		
55	Lys	Trp	Pro	Thr	Ser	Gly	Gln	Leu	Phe	Leu	Gly	His	Pro	Lys	Ala	Ile	
				405				410					415				
	Leu	Gly	Ala	Thr	Glu	Val	Lys	Leu	Leu	Gly	His	Gly	Gln	Pro	Leu	Asn	
60				420				425					430				

280

Trp Ile Ser Leu Glu Gln Asn Gly Ile Met Val Glu Leu Pro Gln Leu
 435 440 445
 5 Thr Ile His Gln Met Pro Cys Lys Trp Gly Trp Ala Leu Ala Leu Thr
 450 455 460
 Asn Val Ile
 465
 10
 (2) INFORMATION FOR SEQ ID NO: 135:
 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 222 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 135:
 20 Met Trp Ser Ala Gly Arg Gly Gly Ala Ala Trp Pro Val Leu Leu Gly
 1 5 10 15
 25 Leu Leu Leu Ala Leu Leu Val Pro Gly Gly Gly Ala Ala Lys Thr Gly
 20 25 30
 30 Ala Glu Leu Val Thr Cys Gly Ser Val Leu Lys Leu Leu Asn Thr His
 35 40 45
 35 His Arg Val Arg Leu His Ser His Asp Ile Lys Tyr Gly Ser Gly Ser
 50 55 60
 Gly Gln Gln Ser Val Thr Gly Val Glu Ala Ser Asp Asp Ala Asn Ser
 65 70 75 80
 40 Tyr Trp Arg Ile Arg Gly Gly Ser Glu Gly Gly Cys Arg Arg Gly Ser
 85 90 95
 Pro Val Arg Cys Gly Gln Ala Val Arg Leu Thr His Val Leu Thr Gly
 100 105 110
 45 Lys Asn Leu His Thr His His Phe Pro Ser Pro Leu Ser Asn Asn Gln
 115 120 125
 Glu Val Ser Ala Phe Gly Glu Asp Gly Glu Gly Asp Asp Leu Asp Leu
 130 135 140
 50 Trp Thr Val Arg Cys Ser Gly Gln His Trp Glu Arg Glu Ala Ala Val
 145 150 155 160
 Arg Phe Gln His Val Gly Thr Ser Val Phe Leu Ser Val Thr Gly Glu
 165 170 175
 55 Gln Tyr Gly Ser Pro Ile Arg Gly Gln His Glu Val His Gly Met Pro
 180 185 190
 Ser Ala Asn Thr His Asn Thr Trp Lys Ala Met Glu Gly Ile Phe Ile
 195 200 205
 60 Lys Pro Ser Val Glu Pro Ser Ala Gly His Asp Glu Leu Xaa

281

210

215

220

5 (2) INFORMATION FOR SEQ ID NO: 136:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 156 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 136:

Met Val Ile Glu Ile Ser Asn Lys Thr Ser Ser Ser Ser Thr Cys Ile
 1 5 10 15
 Leu Val Leu Leu Val Ser Phe Cys Leu Leu Leu Val Pro Ala Met Tyr
 20 25 30
 Ser Ser Asp Thr Arg Gly Ser Leu Pro Ala Glu His Gly Val Leu Ser
 35 40 45
 Arg Gln Leu Arg Ala Leu Pro Ser Glu Asp Pro Tyr Gln Leu Glu Leu
 50 55 60
 Pro Ala Leu Gln Ser Glu Val Pro Lys Asp Ser Thr His Gln Trp Leu
 65 70 75 80
 Asp Gly Ser Asp Cys Val Leu Gln Ala Pro Gly Asn Thr Ser Cys Leu
 85 90 95
 Leu His Tyr Met Pro Gln Ala Pro Ser Ala Glu Pro Pro Leu Glu Trp
 100 105 110
 Pro Phe Pro Asp Leu Phe Ser Glu Pro Leu Cys Arg Gly Pro Ile Leu
 115 120 125
 Pro Leu Gln Ala Asn Leu Thr Arg Lys Gly Gly Trp Leu Pro Thr Gly
 130 135 140
 Ser Pro Ser Val Ile Leu Gln Asp Arg Tyr Ser Gly
 145 150 155

45 (2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 233 amino acids

(B) TYPE: amino acid

50 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 137:

Met Met Ile Leu Phe Asn Leu Leu Ile Phe Leu Cys Gly Ala Ala Leu
 1 5 10 15
 Leu Ala Val Gly Ile Trp Val Ser Ile Asp Gly Ala Ser Phe Leu Lys
 20 25 30
 Ile Phe Gly Pro Leu Ser Ser Ser Ala Met Gln Phe Val Asn Val Gly
 35 40 45

60

282

Tyr Phe Leu Ile Ala Ala Gly Val Val Val Phe Ala Leu Gly Phe Leu
 50 55 60
 5 Gly Cys Tyr Gly Ala Lys Thr Glu Ser Lys Cys Ala Leu Val Thr Phe
 65 70 75 80
 Phe Phe Ile Leu Leu Leu Ile Phe Ile Ala Glu Val Ala Ala Ala Val
 85 90 95
 10 Val Ala Leu Val Tyr Thr Thr Met Ala Glu His Phe Leu Thr Leu Leu
 100 105 110
 Val Val Pro Ala Ile Lys Lys Asp Tyr Gly Ser Gln Glu Asp Phe Thr
 115 120 125
 15 Gln Val Trp Asn Thr Thr Met Lys Gly Leu Lys Cys Cys Gly Phe Thr
 130 135 140
 20 Asn Tyr Thr Asp Phe Glu Asp Ser Pro Tyr Phe Lys Glu Asn Ser Ala
 145 150 155 160
 Phe Pro Pro Phe Cys Cys Asn Asp Asn Val Thr Asn Thr Ala Asn Glu
 165 170 175
 25 Thr Cys Thr Lys Gln Lys Ala His Asp Gln Lys Val Glu Gly Cys Phe
 180 185 190
 30 Asn Gln Leu Leu Tyr Asp Ile Arg Thr Asn Ala Val Thr Val Gly Gly
 195 200 205
 Val Ala Ala Gly Ile Gly Gly Leu Glu Leu Ala Ala Met Ile Val Ser
 210 215 220
 35 Met Tyr Leu Tyr Cys Asn Leu Gln Xaa
 225 230
 40 (2) INFORMATION FOR SEQ ID NO: 138:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 61 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:
 Met Gly Ser Ser Arg Trp Ser Val Ala Cys Pro Thr Gly Leu Gly Val
 1 5 10 15
 50 Leu Met Leu Gly Leu Gly Gly Asp His Pro Pro Gly Ser Gln Val Asp
 20 25 30
 Pro Leu Leu Met Gly Xaa Cys Val Arg Pro Xaa Leu Pro Glu Leu Thr
 35 40 45
 55 Ala Xaa Trp Arg Glu Xaa Gln Xaa Arg Ser Ala Ser Ala
 50 55 60

60

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 139:

5 Met Gly Trp Leu Phe Leu Lys Val Leu Leu Ala Gly Val Ser Phe Ser
 1 5 10 15
 Gly Phe Leu Tyr Pro Leu Val Asp Phe Cys Ile Ser Gly Lys Thr Arg
 20 25 30
 15 Gly Gln Lys Pro Asn Phe Val Ile Ile Leu Ala Asp Asp Met Gly Trp
 35 40 45
 Gly Asp Trp Gly Ala Asn Trp Ala Glu Thr Lys Asp Thr Ala Asn Leu
 20 50 55 60
 Asp Lys Met Ala Ser Glu Gly Met Xaa
 65 70

25

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 377 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

35 Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Leu Met
 1 5 10 15
 Gln Phe Leu Cys His Glu Phe Leu Arg Gly Asn Pro Arg Val Thr Arg
 20 25 30
 40 Leu Leu Ser Glu Met Arg Ile His Leu Leu Pro Ser Met Asn Pro Asp
 35 40 45
 Gly Tyr Glu Ile Ala Tyr His Arg Gly Ser Glu Leu Val Gly Trp Ala
 45 50 55 60
 Glu Gly Arg Trp Asn Asn Gln Ser Ile Asp Leu Asn His Asn Phe Ala
 65 70 75 80
 50 Asp Leu Asn Thr Pro Leu Trp Glu Ala Gln Asp Asp Gly Lys Val Pro
 85 90 95
 His Ile Val Pro Asn His His Leu Pro Leu Pro Thr Tyr Tyr Thr Leu
 100 105 110
 55 Pro Asn Ala Thr Val Ala Pro Glu Thr Arg Ala Val Ile Lys Trp Met
 115 120 125
 60 Lys Arg Ile Pro Phe Val Leu Ser Ala Asn Leu His Gly Gly Glu Leu
 130 135 140

284

Val Val Ser Tyr Pro Phe Asp Met Thr Arg Thr Pro Trp Ala Ala Arg
 145 150 155 160
 5 Glu Leu Thr Pro Thr Pro Asp Asp Ala Val Phe Arg Trp Leu Ser Thr
 165 170 175
 Val Tyr Ala Gly Ser Asn Leu Ala Met Gln Asp Thr Ser Arg Arg Pro
 180 185 190
 10 Cys His Ser Gln Asp Phe Ser Val His Gly Asn Ile Ile Asn Gly Ala
 195 200 205
 Asp Trp His Thr Val Pro Gly Ser Met Asn Asp Phe Ser Tyr Leu His
 210 215 220
 Thr Asn Cys Phe Glu Val Thr Val Glu Leu Ser Cys Asp Lys Phe Pro
 225 230 235 240
 20 His Glu Asn Glu Leu Pro Gln Glu Trp Glu Asn Asn Lys Asp Ala Leu
 245 250 255
 Leu Thr Tyr Leu Glu Gln Val Arg Met Gly Ile Ala Gly Val Val Arg
 260 265 270
 25 Asp Lys Asp Thr Glu Leu Gly Ile Ala Asp Ala Val Ile Ala Val Asp
 275 280 285
 Gly Ile Asn His Asp Val Thr Thr Ala Trp Gly Gly Asp Tyr Trp Arg
 290 295 300
 30 Leu Leu Thr Pro Gly Asp Tyr Met Val Thr Ala Ser Ala Glu Gly Tyr
 305 310 315 320
 35 His Ser Val Thr Arg Asn Cys Arg Val Thr Phe Glu Glu Gly Pro Phe
 325 330 335
 Pro Cys Asn Phe Val Leu Thr Lys Thr Pro Lys Gln Arg Leu Arg Glu
 340 345 350
 40 Leu Leu Ala Ala Gly Ala Lys Val Pro Pro Asp Leu Arg Arg Arg Leu
 355 360 365
 Glu Arg Leu Arg Gly Gln Lys Asp Xaa
 370 375
 45

(2) INFORMATION FOR SEQ ID NO: 141:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

Met Ile Cys Leu Ile Leu Leu Leu Gln Ala Val Val Phe Leu Arg Ser
 1 5 10 15

60 Leu His Val Val His Asn Phe Gln Ile Leu Asp Leu Ser Gly Thr Ser

285

20 25

Tyr Pro Lys Phe Tyr Gln Thr Leu His Arg Gln

35 40

(2) INFORMATION FOR SEQ ID NO: 142:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 41 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

```

15 Met Val His Val Leu Glu Ile Leu Leu Phe Ile Thr Met Gln Ala Val
    1             5             10             15

Ser Phe Pro Phe Gln Thr Gln Ile Asp Thr Cys Asn Thr Gln Asp Pro
    20             25             30

Ala Glu Arg Gln Pro Ala Ser Ile Val
    35             40

```

25

(2) INFORMATION FOR SEQ ID NO: 143:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 143:

```

35 Met Gly Ser Cys Ser Lys Asn Arg Ser Phe Phe Trp Met Thr Gly Leu
    1          5          10          15

Leu Val Phe Ile Ser Leu Leu Leu Ser Glu Trp Gln Gly Pro Trp Glu
    20          25          30

40 Gly Arg Ala Ile Gly Glu Gly Trp Ala Ser Trp Ala Leu Thr Asn Gly
    35          40          45

Trp Ala Val Gln Leu Leu Met Ser Leu Gly Asn Asn Thr Glu Lys His
    50          55          60

45 Ser Val Met Ile Tyr Glu
    65          70

```

50

(2) INFORMATION FOR SEQ ID NO: 144:

55 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 483 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

60 Met Ala Thr Gly Gly Gly Ile Arg Ala Met Thr Ser Leu Tyr Gly Glx.

286

	1	5	10	15
	Leu Ala Gly	Leu Lys Glu	Leu Gly Leu Asp Cys	Xaa Ser Tyr Ile
		20	25	30
5	Thr Gly Ala	Ser Gly Ser Thr Trp	Ala Leu Ala Asn Leu Tyr Lys Asp	
		35	40	45
	Pro Glu Trp	Ser Gln Lys Asp Leu Ala Gly	Pro Thr Glu Leu Leu Lys	
10		50	55	60
	Thr Gln Val	Thr Lys Asn Lys Leu Gly Val	Leu Ala Pro Ser Gln Leu	
		65	70	75
	Gln Arg Tyr	Arg Gln Glu Leu Ala Glu Arg	Ala Arg Leu Gly Tyr Pro	
15		85	90	95
	Ser Cys Phe	Thr Asn Leu Trp Ala Leu Ile	Asn Glu Ala Leu Leu His	
		100	105	110
20	Asp Glu Pro	His Asp His Lys Leu Ser Asp	Gln Arg Glu Ala Leu Ser	
		115	120	125
	His Gly Gln	Asn Pro Leu Pro Ile Tyr Cys	Ala Leu Asn Thr Lys Gly	
25		130	135	140
	Gln Ser Leu	Thr Thr Phe Glu Phe Gly Glu	Trp Cys Glu Phe Ser Pro	
		145	150	155
	Tyr Glu Val	Gly Phe Pro Lys Tyr Gly Ala Phe	Ile Pro Ser Glu Leu	
30		165	170	175
	Phe Gly Ser	Glu Phe Phe Met Gly Gln Leu Met	Lys Arg Leu Pro Glu	
		180	185	190
35	Ser Arg Ile	Cys Phe Leu Glu Gly Ile Trp Ser	Asn Leu Tyr Ala Ala	
		195	200	205
	Asn Leu Gln	Asp Ser Leu Tyr Trp Ala Ser Glu	Pro Ser Gln Phe Trp	
40		210	215	220
	Asp Arg Trp	Val Arg Asn Gln Ala Asn Leu Asp	Lys Glu Gln Val Pro	
		225	230	235
	Leu Leu Lys	Ile Glu Glu Pro Pro Ser Thr Ala	Gly Arg Ile Ala Glu	
45		245	250	255
	Phe Phe Thr	Asp Leu Leu Thr Trp Arg Pro Leu	Ala Gln Ala Thr His	
		260	265	270
50	Asn Phe Leu	Arg Gly Leu His Phe His Lys Asp	Tyr Phe Gln His Pro	
		275	280	285
	His Phe Ser	Thr Trp Lys Ala Thr Thr Leu Asp	Gly Leu Pro Asn Gln	
55		290	295	300
	Leu Thr Pro	Ser Glu Pro His Leu Cys Leu Leu Asp	Val Gly Tyr Leu	
		305	310	315
	Ile Asn Thr	Ser Cys Leu Pro Leu Leu Gln Pro	Thr Arg Asp Val Asp	
60				

287

325 330 335
 Leu Ile Leu Ser Leu Asp Tyr Asn Leu His Gly Ala Phe Gln Gln Leu
 340 345 350
 5 Gln Leu Leu Gly Arg Phe Cys Gln Glu Gln Gly Ile Pro Phe Pro Pro
 355 360 365
 10 Ile Ser Pro Ser Pro Glu Glu Gln Leu Gln Pro Arg Glu Cys His Thr
 370 375 380
 Phe Ser Asp Pro Thr Cys Pro Gly Ala Pro Ala Val Leu His Phe Pro
 385 390 395 400
 15 Leu Val Ser Asp Ser Phe Arg Glu Tyr Ser Ala Pro Gly Val Arg Arg
 405 410 415
 Thr Pro Glu Glu Ala Ala Ala Gly Glu Val Asn Leu Ser Ser Ser Asp
 420 425 430
 20 Ser Pro Tyr His Tyr Thr Lys Val Thr Tyr Ser Gln Glu Asp Val Asp
 435 440 445
 Lys Leu Leu His Leu Thr His Tyr Asn Val Cys Asn Asn Gln Glu Gln
 450 455 460
 25 Leu Leu Glu Ala Leu Arg Gln Ala Val Gln Arg Arg Arg Gln Arg Arg
 465 470 475 480
 30 Pro His Xaa

35 (2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 226 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 145:

Met Glu Gly Ala Pro Pro Gly Ser Leu Ala Leu Arg Leu Leu Leu Phe
 1 5 10 15
 45 Val Ala Leu Pro Ala Ser Gly Trp Leu Thr Thr Gly Ala Pro Glu Pro
 20 25 30
 Pro Pro Leu Ser Gly Ala Pro Gln Asp Gly Ile Arg Ile Asn Val Thr
 35 40 45
 50 Thr Leu Lys Asp Asp Gly Asp Ile Ser Lys Gln Gln Val Val Leu Asn
 50 55 60
 55 Ile Thr Tyr Glu Ser Gly Gln Val Tyr Val Asn Asp Leu Pro Val Asp
 65 70 75 80
 Ser Gly Val Thr Arg Ile Ser Cys Gln Thr Leu Ile Val Lys Asn Glu
 85 90 95
 60

288

Asn Leu Glu Asn Leu Glu Glu Lys Glu Tyr Phe Gly Ile Val Ser Val
 100 105 110
 5 Arg Ile Leu Val His Glu Trp Pro Met Thr Ser Gly Ser Ser Leu Gln
 115 120 125
 Leu Ile Val Ile Gln Glu Glu Val Val Glu Ile Asp Gly Lys Gln Val
 130 135 140
 10 Gln Gln Lys Asp Val Thr Glu Ile Asp Ile Leu Val Lys Asn Arg Gly
 145 150 155 160
 Val Leu Arg His Ser Asn Tyr Thr Leu Pro Leu Glu Glu Ser Met Leu
 165 170 175
 15 Tyr Ser Ile Ser Arg Asp Ser Asp Ile Leu Phe Thr Leu Pro Asn Leu
 180 185 190
 20 Ser Lys Lys Glu Ser Val Ser Ser Leu Gln Thr Thr Ser Gln Tyr Leu
 195 200 205
 Ile Arg Asn Val Glu Thr Thr Val Asp Glu Asp Val Leu Pro Gly Gln
 210 215 220
 25 Val Thr
 225

30 (2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 amino acids

(B) TYPE: amino acid

35 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 146:

Met Gly Met Gly Ala Phe Gln Ala Phe Phe Trp Val Ile Leu Thr Val
 1 5 10 15
 40 Ser Asn Val Cys Val Leu Phe Lys Met Ser Leu Phe Phe Leu Leu Thr
 20 25 30
 45 Leu Ile Ser Lys Leu His Gly Asp Ala Glu Val Cys Xaa
 35 40 45

50 (2) INFORMATION FOR SEQ ID NO: 147:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 132 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 147:

Met Ser Gly Gly Trp Met Ala Gln Val Gly Ala Trp Arg Thr Gly Ala
 1 5 10 15
 60 Leu Gly Leu Ala Leu Leu Leu Leu Gly Leu Gly Leu Gly Leu Glu

289

20 25 30

Ala Pro Arg Ala Arg Phe Pro Pro Arg Pro Leu Pro Arg Pro His Pro
35 40 45

5 Ser Ser Gly Ser Cys Pro Pro Thr Lys Phe Gln Cys Arg Thr Ser Gly
50 55 60

10 Leu Cys Val Pro Leu Thr Trp Arg Cys Asp Arg Thr Trp Thr Ala Ala
65 70 75 80

Met Ala Ala Met Arg Arg Ser Ala Gly Leu Ser His Val Pro Arg Lys
85 90 95

15 Gly Asn Ala His Arg Pro Leu Ala Ser Pro Ala Pro Ala Pro Ala Ser
100 105 110

Val Thr Ala Leu Gly Glu Leu Thr Arg Asn Cys Ala Thr Ala Ala Ala
115 120 125

20 Trp Pro Ala Xaa
130

25

(2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 92 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 148:

35 Met Glu Ala Thr Leu Glu Gln His Leu Glu Asp Thr Met Lys Asn Pro
1 5 10 15

Ser Ile Val Gly Val Leu Cys Thr Asp Ser Gln Gly Leu Asn Leu Gly
20 25 30

40 Cys Arg Gly Thr Leu Ser Asp Glu His Ala Gly Val Ile Ser Val Leu
35 40 45

Ala Gln Gln Ala Ala Lys Leu Thr Ser Asp Pro Thr Asp Ile Pro Val
50 55 60

45 Val Cys Leu Glu Ser Asp Asn Gly Asn Ile Met Ile Gln Lys His Asp
65 70 75 80

50 Gly Ile Thr Val Ala Val His Lys Met Ala Ser Xaa
85 90

55

(2) INFORMATION FOR SEQ ID NO: 149:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 165 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

290

Met Glu Pro Leu Arg Leu Leu Ile Leu Leu Phe Val Thr Glu Leu Ser
1 5 10 15

5 Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu
20 25 30

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys
35 40 45

10 Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val
50 55 60

Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly
15 65 70 75 80

Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr
85 90 95

20 Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser
100 105 110

Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val
115 120 125

25 Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro
130 135 140

Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser
145 150 155 160

30 Arg Ser Ser Ser Xaa
165

35

(2) INFORMATION FOR SEQ ID NO: 150:

- 40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 150:

45 Met Ile Ser Leu Thr Asp Thr Gln Lys Ile Gly Met Gly Leu Thr Gly
1 5 10 15

Phe Gly Val Phe Phe Leu Phe Phe Gly Met Ile Leu Phe Phe Asp Lys
20 25 30

50 Ala Leu Leu Ala Ile Gly Asn Val Leu Phe Val Ala Gly Leu Ala Phe
35 40 45

Val Ile Gly Leu Glu Arg Thr Phe Arg Phe Phe Phe Gln Lys His Lys
55 60

Met Lys Ala Thr Gly Phe Phe Leu Gly Gly Val Phe Val Val Leu Ile
65 70 75 80

60 Gly Trp Pro Leu Ile Gly Met Ile Phe Glu Ile Tyr Gly Phe Phe Leu

291

85 90 95

Leu Phe Arg Gly Phe Phe Pro Val Val Val Gly Phe Ile Arg Arg Val
100 105 110

5 Pro Val Leu Gly Ser Leu Leu Asn Leu Pro Gly Ile Arg Ser Phe Val
115 120 125

10 Asp Lys Val Gly Glu Ser Asn Asn Met Val Xaa
130 135

15 (2) INFORMATION FOR SEQ ID NO: 151:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 151:

Met Ser Ala Pro Gln Thr Arg Ile Ser Arg Ala Leu Val Leu Leu Phe
1 5 10 15

25 Leu Ala Pro Thr Leu Leu Ser Leu Gly His Gly Ile His Pro Ile Asn
20 25 30

Thr Ala Thr Pro Tyr Xaa Thr Asp Gln Ala Lys Leu Ala Pro Gly Thr
35 40 45

30 Lys Glu Leu Asn His Asp Gln Ser Val Thr
50 55

35

(2) INFORMATION FOR SEQ ID NO: 152:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 amino acids

40 (B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

Met Ile Arg Lys Leu His Lys Ile Ile Val Phe Ser Pro Arg Val Ile
1 5 10 15

Val Leu Leu Asn Cys Phe Phe Phe Ile Lys Ala Lys Phe Val Leu Tyr
20 25 30

50 Ile Phe Val Phe His Val Leu Asp Gly Ser Ile Ser Tyr Pro Val Xaa
35 40 45

55

(2) INFORMATION FOR SEQ ID NO: 153:

60

(i) SEQUENCE CHARACTERISTICS:

292

(A) LENGTH: 42 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 153:

5

Met Leu Leu Asn Gln His Phe Lys Ile Phe Gly Ser Leu Ile His Met
 1 5 10 15

10

Asn Leu Leu Phe Ala Leu Ile Ser Leu Gly Ser Ser Asn Leu Ser Gly
 20 25 30

Val Gln Phe Cys Cys Glu Thr Val Gln Xaa
 35 40

15

(2) INFORMATION FOR SEQ ID NO: 154:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 72 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

20

25

Met Leu Ser Leu Ser Phe Leu Leu Arg Arg Val Leu Phe Leu Gly Phe
 1 5 10 15

Leu Gln Ala Ser Val Gly Glu Lys Lys Ser Leu Arg Xaa Leu Asn Tyr
 20 25 30

30

Ser Val Pro His Pro Met Leu Xaa His Pro Pro Pro Asp Thr Ala Gln
 35 40 45

35

Val Pro Pro Arg Leu Glu Arg Ser Leu Leu Gln Gln Glu Leu Trp Thr
 50 55 60

Pro Gly Pro His His Ser Asn Ile
 65 70

40

(2) INFORMATION FOR SEQ ID NO: 155:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

45

50

Met Gln Pro Leu Asn Phe Ser Ser Thr Glu Cys Ser Ser Phe Ser Pro
 1 5 10 15

Pro Thr Thr Val Ile Leu Leu Ile Leu Leu Cys Phe Glu Gly Leu Leu
 20 25 30

55

Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile
 35 40 45

60

Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg
 50 55 60

293

Trp Ala Lys Lys Thr Lys Trp Met Asn Met Lys Ala Val Phe Gly His
 65 70 75 80

5 Pro Phe Ser Leu Gly Trp Ala Ser Pro Phe Ala Thr Pro Asp Gln Gly
 85 90 95

Lys Ala Asp Pro Tyr Gln Tyr Val Val Xaa
 100 105

10

(2) INFORMATION FOR SEQ ID NO: 156:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:

20 Met Tyr Thr Asn His Phe Asn Leu Tyr Leu Lys Tyr Ile Leu Leu Ile
 1 5 10 15

Ile Leu Ile Leu Asn Met Thr Asn Ser Ser Ser Arg Tyr
 25 20 25

(2) INFORMATION FOR SEQ ID NO: 157:

30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 53 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:

Met Asn Glu Leu Leu Leu Phe Phe Phe Phe Phe Phe Phe Thr Phe
 1 5 10 15

40 Cys Ile Glu Thr Asn Ser Phe Lys Gln Thr Tyr Tyr Tyr Tyr Phe Leu
 20 25 30

Gln Asn Ile Tyr Met Glu Met Leu Pro Pro Pro Val Asn Pro Pro Val
 35 40 45

45 Pro Pro Trp Gly Xaa
 50

50

(2) INFORMATION FOR SEQ ID NO: 158:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 75 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 158:

60 Met Tyr Ala Val Tyr Gln Gln Leu Ala Gln Leu Thr Leu Met Val Thr
 1 5 10 15

294

Leu Leu Ala Pro Ile Leu Pro Asp Glu Gln Ser Glu Val Phe Glu Ala
 20 25 30
 5 Leu Ser Asn Leu Pro Lys Val Thr Trp Leu Gly Ser Asn Ser Pro Ser
 35 40 45
 Ser Glu Met Pro Glu Pro Gly Arg Phe Val Ile Val His His Gln Leu
 50 55 60
 10 Ser Ala Ala Ser His Ser Ser Ser Gln Leu Ala
 65 70 75

15

(2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 81 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 159:

25 Met Trp Pro Pro Leu Leu Leu Leu Leu Leu Leu Pro Ala Ala Pro
 1 5 10 15
 Val Pro Thr Ala Lys Ala Ala Pro His Pro Asp Ala Asn Thr Gln Glu
 20 25 30
 30 Gly Leu Gln Asn Leu Leu Gln Gly Val Gly Ala Gly Gly Asp Gly Glu
 35 40 45
 Leu Arg Ala Asp Ser His Leu Ala Pro Gly Ser Gly Cys Ile Asp Gly
 50 55 60
 35 Ala Val Val Ala Thr Arg Pro Glu Ser Arg Gly Gly Arg Pro Ala Val
 65 70 75 80
 40 Pro

45

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

50 Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu
 5 10 15
 55 Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala
 20 25 30
 Gln Glu Ala Gly Thr Ser Lys Pro Asn Glu Glu Ile Ser Gly Pro Ala
 35 40 45
 60

295

Glu Pro Ala Ser Pro Pro Glu Thr Thr Thr Thr Ala Gln Glu Thr Ser
 50 55 60
 Ala Ala Ala Val Gln Gly Thr Ala Lys Val Thr Ser Ser Arg Gln Glu
 5 65 70 75 80
 Leu Asn Pro Leu Lys Ser Ile Val Glu Lys Ser Ile Leu Leu Thr Glu
 85 90 95
 10 Gln Ala Leu Ala Lys Ala Gly Lys Gly Met His Gly Gly Val Pro Gly
 100 105 110
 Gly Lys Gln Phe Ile Glu Asn Gly Ser Glu Phe Ala Gln Lys Leu Leu
 115 120 125
 15 Lys Lys Phe Ser Leu Leu Lys Pro Trp Ala Xaa
 130 135
 20 (2) INFORMATION FOR SEQ ID NO: 161:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 178 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:
 Met Leu Gly Cys Gly Ile Pro Ala Leu Gly Leu Leu Leu Leu Gln
 30 1 5 10 15
 Gly Ser Ala Asp Gly Asn Gly Ile Gln Gly Phe Phe Tyr Pro Trp Ser
 20 25 30
 35 Cys Glu Gly Asp Ile Trp Asp Arg Glu Ser Cys Gly Gly Gln Ala Ala
 35 40 45
 Ile Asp Ser Pro Asn Leu Cys Leu Arg Leu Arg Cys Cys Tyr Arg Asn
 50 55 60
 40 Gly Val Cys Tyr His Gln Arg Pro Asp Glu Asn Val Arg Arg Lys His
 65 70 75 80
 Met Trp Ala Leu Val Trp Thr Cys Ser Gly Leu Leu Leu Leu Ser Cys
 45 85 90 95
 Ser Ile Cys Leu Phe Trp Trp Ala Lys Arg Arg Asp Val Leu His Met
 100 105 110
 50 Pro Gly Phe Leu Ala Gly Pro Cys Asp Met Ser Lys Ser Val Ser Leu
 115 120 125
 Leu Ser Lys His Arg Gly Thr Lys Lys Thr Pro Ser Thr Gly Ser Val
 130 135 140
 55 Pro Val Ala Leu Ser Lys Glu Ser Arg Asp Val Glu Gly Gly Thr Glu
 145 150 155 160
 Gly Glu Gly Thr Glu Glu Gly Glu Glu Thr Glu Gly Glu Glu Glu
 60 165 170 175

296

Asp Xaa

5

(2) INFORMATION FOR SEQ ID NO: 162:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 72 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 162:

15

Met Glu Ala Val Phe Thr Val Phe Phe Phe Val Val Val Leu Phe Leu
1 10 15Lys Asn Thr Glu Gly Ala Lys Leu Phe Cys Thr Leu Tyr Pro Ala Ala
20 25 30

20

Ser Ser Gly Gln Ser Gln Gly Pro Gly Leu Glu Lys Pro Asp Ser Gln
35 40 45

25

Glu Cys Ile Ile Asp Pro Cys Ser Tyr Pro Ile Ala Leu Gly Ala Gly
50 55 60Thr Glu Pro Gly Cys Lys Ile Xaa
65 70

30

(2) INFORMATION FOR SEQ ID NO: 163:

(i) SEQUENCE CHARACTERISTICS:

35

(A) LENGTH: 67 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 163:

40

Met Trp Phe Tyr Phe Leu Ser Val Ser Phe Pro Leu Leu Pro Val Xaa
1 10 15Ala Pro Xaa Pro Pro Pro Ala Pro Thr Thr Leu Cys Leu Leu Leu Phe
20 25 30

45

Leu Gly Xaa Leu Tyr Asn Ser Thr Cys Ile His Cys Val His Thr Thr
35 40 45

50

Ser Xaa Thr Gln Asn Pro Thr Ala Asn Thr Leu Lys Lys Lys Lys Lys
50 55 60Asn Trp Gly
65

55

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

60

(A) LENGTH: 155 amino acids

297

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:

5 Met Gly Phe Gly Ala Thr Leu Ala Val Gly Leu Thr Ile Phe Val Leu
 1 5 10 15
 Ser Val Val Thr Ile Ile Ile Cys Phe Thr Cys Ser Cys Cys Cys Leu
 20 25 30
 10 Tyr Lys Thr Cys Arg Arg Pro Arg Pro Val Val Thr Thr Thr Ser
 35 40 45
 Thr Thr Val Val His Ala Pro Tyr Pro Gln Pro Pro Ser Val Pro Pro
 15 50 55 60
 Ser Tyr Pro Gly Pro Ser Tyr Gln Gly Tyr His Thr Met Pro Pro Gln
 65 70 75 80
 20 Pro Gly Met Pro Ala Ala Pro Tyr Pro Met Gln Tyr Pro Pro Pro Tyr
 85 90 95
 Pro Ala Gln Pro Met Gly Pro Pro Ala Tyr His Glu Thr Leu Ala Gly
 100 105 110
 25 Glu Gln Pro Arg Pro Thr Pro Pro Ala Ser Leu Leu Thr Thr Arg Pro
 115 120 125
 Thr Trp Met Pro Arg Arg Arg Pro Ser Glu His Ser Leu Ala Ser Leu
 130 135 140
 30 Ala Ala Thr Trp Leu Cys Cys Val Cys Ala Xaa
 145 150 155

35

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:

40

(A) LENGTH: 104 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

45 Met Ile Ile Leu Val Phe Ile Ala Phe Phe Ile Pro Leu Gln Lys Thr
 1 5 10 15
 Ile Gly Lys Ile Ala Thr Cys Leu Glu Leu Arg Ser Ala Ala Leu Gln
 20 25 30
 50 Ser Thr Gln Ser Gln Glu Glu Phe Lys Leu Glu Asp Leu Lys Lys Leu
 35 40 45
 Glu Pro Ile Leu Lys Asn Ile Leu Thr Tyr Asn Lys Glu Phe Pro Phe
 55 60
 Asp Val Gln Pro Val Pro Leu Arg Arg Ile Leu Ala Pro Gly Glu Glu
 65 70 75 80
 60 Glu Asn Leu Glu Phe Glu Glu Asp Glu Glu Glu Gly Gly Ala Gly Ala

298

88

90

92

Gly Leu Leu Ile Leu Ser Cys Xaa
100

5

(2) INFORMATION FOR SEQ ID NO: 166:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 81 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

15

Met Ala Gly Thr Met Val Ile Val Val Val Val Val Val Gly Glu Val
1 5 10 15

20

Val Val Glu Ala Glu Val Val Val Gln Ala Arg Glu Glu Ala Gly Glu
20 25 30

Glu Glu Gly Ala Arg Ile Ile Thr Lys Gly Val Asn Leu Asn Ser Ile
35 40 45

25

Ser Ser Met Glu Val Ile Ser Ile Ile Ile Leu Asp Leu Asp Arg Glu
50 55 60

Asp Ile Thr Leu Val Glu Ala Thr Glu Pro Tyr Ile Leu Leu Glu Leu
65 70 75 80

30

Lys

35

(2) INFORMATION FOR SEQ ID NO: 167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 93 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

40

Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
1 5 10 15

Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys
20 25 30

50

Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val
35 40 45

Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu
50 55 60

55

Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser
65 70 75 80

Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr
85 90

60

(2) INFORMATION FOR SEQ ID NO: 168:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

Met Gly Trp Ser Ala Gly Leu Leu Phe Leu Leu Ile Leu Tyr Leu Pro
1 5 10 15

15 Val Pro Gly Trp Met Glu Arg Glu Asp Gly Glu Thr Gly His Leu Ser
20 25 30

Pro Gln Ala Pro Gly Arg Glu Tyr Arg Gly Phe Tyr Ser Val Pro Pro
35 40 45

20

Asp Tyr Val Trp Leu Arg Asp Ser Pro Xaa
50 55

25

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 232 amino acids

30

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 169:

35 Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
1 5 10 15

Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Ala His Cys Gln Thr
20 25 30

40 Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
35 40 45

Ile Lys Lys Asn Ser Gly His Ile Tyr Asn Lys Asn Ile Ser Gln Lys
50 55 60

45

Asp Cys Asp Cys Leu His Val Val Glu Pro Met Pro Val Arg Gly Pro
65 70 75 80

50 Asp Val Glu Ala Tyr Cys Leu Arg Cys Glu Cys Lys Tyr Glu Glu Arg
85 90 95

Ser Ser Val Thr Ile Lys Val Thr Ile Ile Ile Tyr Leu Ser Ile Leu
100 105 110

55 Gly Leu Leu Leu Leu Tyr Met Val Tyr Leu Thr Leu Val Glu Pro Ile
115 120 125

Leu Lys Arg Arg Leu Phe Gly His Ala Gln Leu Ile Gln Ser Asp Asp
130 135 140

60

300

Asp Ile Gly Asp His Gln Pro Phe Ala Asn Ala His Asp Val Leu Ala
 145 150 155 160
 5 Arg Ser Arg Ser Arg Ala Asn Val Leu Asn Lys Val Glu Tyr Gly Trp
 165 170 175
 Ala Ala Leu Glu Ala Ser Ser Pro Arg Ala Ala Lys Ser Leu Ser Leu
 180 185 190
 10 Thr Gly Met Leu Ser Ser Ala Asn Trp Gly Ile Glu Phe Lys Val Thr
 195 200 205
 Arg Lys Lys Gln Ala Asp Asn Trp Lys Gly Thr Asp Trp Val Leu Leu
 210 215 220
 15 Gly Phe Ile Leu Ile Pro Cys Xaa
 225 230

20

(2) INFORMATION FOR SEQ ID NO: 170:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 72 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

30 Met Ser Ala Ile Phe Asn Phe Gln Ser Leu Leu Thr Val Ile Leu Leu
 1 5 10 15
 Leu Ile Cys Thr Cys Ala Tyr Ile Arg Ser Leu Ala Pro Ser Leu Leu
 20 25 30
 35 Asp Arg Asn Lys Thr Gly Leu Leu Gly Ile Phe Trp Lys Cys Ala Arg
 35 40 45
 Ile Gly Glu Arg Lys Ser Pro Tyr Val Ala Val Cys Cys Ile Val Met
 50 55 60
 40 Ala Phe Ser Ile Leu Phe Ile Gln
 65 70

45

(2) INFORMATION FOR SEQ ID NO: 171:

(i) SEQUENCE CHARACTERISTICS:

50

(A) LENGTH: 65 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

55 Met Gly Thr Phe Ser Leu Ser Leu Phe Gly Leu Met Gly Val Ala Phe
 1 5 10 15
 Gly Met Asn Leu Glu Ser Ser Leu Glu Glu Asp His Arg Ile Phe Trp
 20 25 30
 60 Leu Ile Thr Gly Ile Met Phe Met Gly Ser Gly Leu Ile Trp Arg Arg

35 40 45

Leu Leu Ser Phe Leu Gly Arg Gln Leu Glu Ala Pro Leu Pro Pro Met
50 55 60

Val
65

10

(2) INFORMATION FOR SEQ ID NO: 172:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
15 (B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

20 Met Tyr Lys Gly Lys Leu Val Ile Val Leu Ile Leu Leu Leu Leu Pro
1 5 10 15

Ser His Phe Met Phe Leu Thr Gln Cys Lys Glu Ile Lys His Asn Leu
20 25 30

25 Lys Lys Asn Met Ser Leu Leu Leu Phe Thr Ile Lys Ser Trp Leu Tyr
35 40 45

Ser Ala Ser Leu Gly Ile Leu Tyr Asn Trp Gln His Leu Thr Ala Gln
50 55 60

30 Val Asp Gln Cys Thr Ser Leu Ile Leu Ile His
65 70 75

35

(2) INFORMATION FOR SEQ ID NO: 173:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 334 amino acids
40 (B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

45 Met Val Gly His Glu Met Ala Ser Xaa Ser Ser Asn Thr Ser Leu Pro
1 5 10 15

Phe Ser Asn Met Gly Asn Pro Met Asn Thr Thr Gln Leu Gly Lys Ser
20 25 30

50 Leu Phe Gln Trp Gln Val Glu Gln Glu Glu Ser Lys Leu Ala Asn Ile
35 40 45

Ser Gln Asp Gln Phe Leu Ser Lys Asp Ala Asp Gly Asp Thr Phe Leu
50 55 60

55 His Ile Ala Val Ala Gln Gly Arg Arg Ala Leu Ser Tyr Val Leu Ala
65 70 75 80

Arg Lys Met Asn Ala Leu His Met Leu Asp Ile Lys Glu His Asn Gly
85 90 95

Gln Ser Ala Phe Gln Val Ala Val Ala Ala Asn Gln His Leu Ile Val
 100 105 110
 5 Gln Asp Leu Val Asn Ile Gly Ala Gln Val Asn Thr Thr Asp Cys Trp
 115 120 125
 Gly Arg Thr Pro Leu His Val Cys Ala Glu Lys Gly His Ser Gln Val
 130 135 140
 10 Leu Gln Ala Ile Gln Lys Gly Ala Val Gly Ser Asn Gln Phe Val Asp
 145 150 155 160
 Leu Glu Ala Thr Asn Tyr Asp Gly Leu Thr Pro Leu His Cys Ala Val
 165 170 175
 15 Ile Ala His Asn Ala Val Val His Glu Leu Gln Arg Asn Gln Gln Pro
 180 185 190
 20 His Ser Pro Glu Val Gln Glu Leu Leu Lys Asn Lys Ser Leu Val
 195 200 205
 Asp Thr Ile Lys Cys Leu Ile Gln Met Gly Ala Ala Val Glu Ala Lys
 210 215 220
 25 Asp Arg Lys Ser Gly Arg Thr Ala Leu His Leu Ala Ala Glu Glu Ala
 225 230 235 240
 Asn Leu Glu Leu Ile Arg Leu Phe Leu Glu Leu Pro Ser Cys Leu Ser
 245 250 255
 30 Phe Val Asn Ala Lys Ala Tyr Asn Gly Asn Thr Ala Leu His Val Ala
 260 265 270
 35 Ala Ser Leu Gln Tyr Arg Leu Thr Gln Leu Asp Ala Val Arg Leu Leu
 275 280 285
 Met Arg Lys Gly Ala Asp Pro Ser Thr Arg Asn Leu Glu Asn Glu Gln
 290 295 300
 40 Pro Val His Leu Val Pro Asp Gly Pro Val Gly Glu Gln Ile Arg Arg
 305 310 315 320
 Ile Leu Lys Gly Lys Ser Ile Gln Gln Arg Ala Pro Pro Tyr
 325 330
 45
 (2) INFORMATION FOR SEQ ID NO: 174:
 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 196 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:
 Met Asp Ala Arg Trp Trp Ala Val Val Val Leu Ala Ala Phe Pro Ser
 1 10 15
 60 Leu Gly Ala Gly Gly Glu Thr Pro Glu Ala Pro Pro Glu Ser Trp Thr

305

[illegible]

(2) INFORMATION FOR SEQ ID NO: 175:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 265 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

45 Met Ser Asp Leu Leu Leu Leu Gly Leu Ile Gly Gly Leu Thr Leu Leu
1 5 10 15

50 Leu Leu Leu Thr Leu Leu Ala Phe Ala Gly Tyr Ser Gly Leu Leu Ala
20 25 30

Gly Val Glu Val Ser Ala Gly Ser Pro Pro Ile Arg Asn Val Thr Val
35 40 45

55 Ala Tyr Lys Phe His Met Gly Leu Tyr Gly Glu Thr Gly Arg Leu Phe
50 55 60

Thr Glu Ser Cys Ser Ile Ser Pro Lys Leu Arg Ser Ile Ala Val Tyr
65 70 75 80

60

304

Tyr Asp Asn Pro His Met Val Pro Pro Asp Lys Cys Arg Cys Ala Val
 85 90 95
 5 Gly Ser Ile Leu Ser Glu Gly Glu Glu Ser Pro Ser Pro Glu Leu Ile
 100 105 110
 Asp Leu Tyr Gln Lys Phe Gly Phe Lys Val Phe Ser Phe Pro Glu Pro
 115 120 125
 10 Ser His Val Val Thr Ala Thr Phe Pro Leu Thr Pro Pro Phe Cys Pro
 130 135 140
 Ile Trp Leu Gly Tyr Pro Pro Cys Pro Ser Cys Leu Gly His Leu His
 145 150 155 160
 15 Gln Gly Ala Glu Ala Val Cys Leu Ser Ser Ala Gly Asp Leu Pro Gly
 165 170 175
 Arg Pro Glu Ser Ile Ser Cys Ala His Trp His Gly Gln Gly Asp Phe
 180 185 190
 Tyr Val Pro Glu Met Lys Glu Thr Glu Trp Lys Trp Arg Gly Leu Val
 195 200 205
 25 Glu Ala Ile Asp Thr Gln Val Asp Gly Thr Gly Ala Asp Thr Met Ser
 210 215 220
 Asp Thr Ser Ser Val Ser Leu Glu Val Ser Pro Gly Ser Arg Glu Thr
 225 230 235 240
 30 Ser Ala Ala Thr Leu Ser Pro Gly Ala Ser Ser Arg Gly Trp Asp Asp
 245 250 255
 Gly Asp Thr Arg Ser Glu His Ser Xaa
 260 265
 35

40 (2) INFORMATION FOR SEQ ID NO: 176:
 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

Met Ala Gln Leu Phe Leu Pro Leu Leu Ala Ala Leu Val Leu Ala Gln
 1 5 10 15
 50 Ala Pro Ala Ala Leu Ala Asp Val Leu Glu Gly Asp Ser Ser Glu Asp
 20 25 30
 Arg Ala Phe Arg Val Arg Ile Ala Gly Asp Ala Pro Leu Gln Gly Val
 35 40 45
 55 Leu Gly Gly Ala Leu Thr Ile Pro Cys His Val His Tyr Leu Arg Pro
 50 55 60
 60 Pro Pro Ser Arg Arg Ala Val Leu Gly Ser Pro Arg Val Lys Trp Thr
 65 70 75 80

305

Phe Leu Ser Arg Gly Arg Glu Ala Glu Val Leu Val Ala Arg Gly Val
85 90 95

5 Arg Val Lys Val Asn Glu Ala Tyr Arg Phe Arg Val Ala Leu Pro Ala
100 105 110

Tyr Pro Ala Ser Leu Thr Asp Val Ser Pro Gly Ala Glu Arg Ala Ala
115 120 125

10 Pro Gln Arg Leu Arg Tyr Leu Ser Leu Xaa
130 135

15

(2) INFORMATION FOR SEQ ID NO: 177:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 179 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:

25 Met Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp
1 5 10 15

Leu Cys Cys Ala Thr Pro Ala His Ala Leu Gln Cys Arg Asp Gly Tyr
20 25 30

30 Glu Pro Cys Val Asn Glu Gly Met Cys Val Thr Tyr His Asn Gly Thr
35 40 45

Gly Tyr Cys Lys Gly Pro Glu Gly Phe Leu Gly Glu Tyr Cys Gln His
50 55 60

35 Arg Asp Pro Cys Glu Lys Asn Arg Cys Gln Asn Gly Gly Thr Cys Val
65 70 75 80

40 Ala Gln Ala Met Leu Gly Lys Ala Thr Cys Arg Cys Ala Ser Gly Phe
85 90 95

Thr Gly Glu Asp Cys Gln Tyr Ser Thr Ser His Pro Cys Phe Val Ser
100 105 110

45 Arg Pro Cys Leu Asn Gly Gly Thr Cys His Met Leu Ser Arg Asp Thr
115 120 125

Tyr Glu Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys Gln Trp
130 135 140

50 Thr Asp Ala Cys Leu Ser His Pro Cys Ala Asn Gly Ser Thr Cys Thr
145 150 155 160

55 Thr Val Ala Asn His Phe Leu Gln Met Pro His Arg Leu His Arg Ala
165 170 175

Glu Val Xaa

60

(1) INFORMATION FOR SEQ ID NO: 178:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 155 amino acids
(E) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

10 Met Thr Arg Gly Gly Pro Gly Gly Arg Pro Gly Leu Pro Gln Pro Pro
1 5 10 15
Pro Leu Leu Leu Leu Leu Leu Pro Leu Leu Leu Val Thr Ala Glu
20 25 30
15 Pro Pro Lys Pro Ala Gly Val Tyr Tyr Ala Thr Ala Tyr Trp Met Pro
35 40 45
Ala Glu Lys Thr Val Gln Val Lys Asn Val Met Asp Lys Asn Gly Asp
20 50 55 60
Ala Tyr Gly Phe Tyr Asn Asn Ser Val Lys Thr Thr Gly Trp Gly Ile
65 70 75 80
25 Leu Glu Ile Arg Ala Gly Tyr Gly Ser Gln Thr Leu Ser Asn Glu Ile
85 90 95
Ile Met Phe Val Ala Gly Phe Leu Glu Gly Tyr Leu Ile Ala Pro His
100 105 110
30 Met Asn Asp His Tyr Thr Asn Leu Tyr Pro Gln Leu Ile Thr Lys Pro
115 120 125
Ser Ile Met Asp Lys Val Gln Asp Phe Met Glu Lys Gln Asp Lys Val
35 130 135 140
Asp Pro Glu Lys Tyr Gln Arg Ile Gln Asp Xaa
145 150 155

40

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 295 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

50 Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His
1 5 10 15
Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp
20 25 30
55 Phe Ser Tyr Lys Arg Xaa Asn Cys Lys Pro Ile Pro Val Asn Leu Gln
35 40 45
Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu
60 55 60

307

Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile
 65 70 75 80
 5 Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys
 85 90 95
 Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln
 100 105 110
 10 Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val
 115 120 125
 Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg
 130 135 140
 Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His
 145 150 155 160
 20 Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys
 165 170 175
 Asn Lys Asn Asp Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn
 180 185 190
 25 Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg
 195 200 205
 Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu
 210 215 220
 Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys
 225 230 235 240
 35 Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro
 245 250 255
 Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser
 260 265 270
 40 Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg
 275 280 285
 Ser Ile Arg Lys Leu Gln Cys
 290 295
 45

(2) INFORMATION FOR SEQ ID NO: 180:

50

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 256 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

Met Arg Pro Ala Ala Leu Arg Gly Ala Leu Leu Gly Cys Leu Cys Leu
 1 5 10 15
 60 Ala Leu Leu Cys Leu Gly Gly Ala Asp Lys Arg Leu Arg Asp Asn His

308

20 25 30
 Glu Trp Lys Lys Leu Ile Met Val Gln His Trp Pro Glu Thr Val Cys
 35 40 45
 5 Glu Lys Ile Gln Asn Asp Cys Arg Asp Pro Pro Asp Tyr Trp Thr Ile
 50 55 60
 10 His Gly Leu Trp Pro Asp Lys Ser Glu Gly Cys Asn Arg Ser Trp Pro
 65 70 75 80
 Phe Asn Leu Glu Glu Ile Lys Asp Leu Leu Pro Glu Met Arg Ala Tyr
 85 90 95
 15 Trp Pro Asp Val Ile His Ser Phe Pro Asn Arg Ser Arg Phe Trp Lys
 100 105 110
 His Glu Trp Glu Lys His Gly Thr Cys Ala Ala Gln Val Asp Ala Leu
 115 120 125
 20 Asn Ser Gln Lys Lys Tyr Phe Gly Arg Ser Leu Glu Leu Tyr Arg Glu
 130 135 140
 Leu Asp Leu Asn Ser Val Leu Leu Lys Leu Gly Ile Lys Pro Ser Ile
 145 150 155 160
 Asn Tyr Tyr Gln Val Ala Asp Phe Lys Asp Ala Leu Ala Arg Val Tyr
 165 170 175
 30 Gly Val Ile Pro Lys Ile Gln Cys Leu Pro Pro Ser Gln Asp Glu Glu
 180 185 190
 Val Gln Thr Ile Gly Gln Ile Glu Leu Cys Leu Thr Lys Gln Asp Gln
 195 200 205
 35 Gln Leu Gln Asn Cys Thr Glu Pro Gly Glu Gln Pro Ser Pro Lys Gln
 210 215 220
 Glu Val Trp Leu Ala Asn Gly Ala Ala Glu Ser Arg Gly Leu Arg Val
 225 230 235 240
 Cys Glu Asp Gly Pro Val Phe Tyr Pro Pro Pro Lys Lys Thr Lys His
 245 250 255

45

50 (2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 324 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

Met Ala Pro Leu Leu Leu Gln Leu Ala Val Leu Gly Ala Ala Leu Ala
 1 5 10 15

60

309

Ala Ala Ala Leu Val Leu Ile Ser Ile Val Ala Phe Thr Thr Ala Thr
 20 25 30

5 Lys Met Pro Ala Leu His Arg His Glu Glu Glu Lys Phe Phe Leu Asn
 35 40 45

Ala Lys Gly Gln Lys Glu Thr Leu Pro Ser Ile Trp Asp Ser Pro Thr
 50 55 60

10 Lys Gln Leu Ser Val Val Val Pro Ser Tyr Asn Glu Glu Lys Arg Leu
 65 70 75 80

Pro Val Met Met Asp Glu Ala Leu Ser Tyr Leu Glu Lys Arg Gln Lys
 85 90 95

15 Arg Asp Pro Ala Phe Thr Tyr Glu Val Ile Val Val Asp Asp Gly Ser
 100 105 110

Lys Asp Gln Thr Ser Lys Val Ala Phe Lys Tyr Cys Gln Lys Tyr Gly
 115 120 125

20 Ser Asp Lys Val Arg Val Ile Thr Leu Val Lys Asn Arg Gly Lys Gly
 130 135 140

25 Gly Ala Ile Arg Met Gly Ile Phe Ser Ser Arg Gly Glu Lys Ile Leu
 145 150 155 160

Met Ala Asp Ala Asp Gly Ala Thr Lys Phe Pro Asp Val Glu Lys Leu
 165 170 175

30 Glu Lys Gly Leu Asn Asp Leu Gln Pro Trp Pro Asn Gln Met Ala Ile
 180 185 190

Ala Cys Gly Ser Arg Ala His Leu Glu Lys Glu Ser Ile Ala Gln Arg
 195 200 205

35 Ser Tyr Phe Arg Thr Leu Leu Met Tyr Gly Phe His Phe Leu Val Trp
 210 215 220

40 Phe Leu Cys Val Lys Gly Ile Arg Asp Thr Gln Cys Gly Phe Lys Leu
 225 230 235 240

Phe Thr Arg Glu Ala Ala Ser Arg Thr Phe Ser Ser Leu His Val Glu
 245 250 255

45 Arg Trp Ala Phe Asp Val Glu Leu Leu Tyr Ile Ala Gln Phe Phe Lys
 260 265 270

Ile Pro Ile Ala Glu Ile Ala Val Asn Trp Thr Glu Ile Glu Gly Ser
 275 280 285

50 Lys Leu Val Pro Phe Trp Ser Trp Leu Gln Met Gly Lys Asp Leu Leu
 290 295 300

55 Phe Ile Arg Leu Arg Tyr Leu Thr Gly Ala Trp Arg Leu Glu Gln Thr
 305 310 315 320

Arg Lys Met Asn

60

(2) INFORMATION FOR SEQ ID NO: 182:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 47 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

10 Met Asp Ile Cys Phe Phe His Tyr Val Leu Leu Phe Phe Leu Val Arg
 1 5 10 15

15 Cys Ala Leu Val Val Leu Ile Leu Leu Cys Gln Gly Trp Gly Asn Gly
 20 25 30

Gly Gly Cys Val Gly Arg Val Leu Ile Ile Val Phe Ser Ser Val
 35 40 45

20

(2) INFORMATION FOR SEQ ID NO: 183:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 93 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

30 Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr
 1 5 10 15

Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp
 20 25 30

35 Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe
 35 40 45

Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe
 40 50 55 60

Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Glu Gly Thr Phe
 65 70 75 80

45 Arg Ser Ser Ile Arg Arg Leu Ser Thr Arg Arg Arg Xaa
 85 90

50 (2) INFORMATION FOR SEQ ID NO: 184:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 168 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

55 Met Xaa Thr Lys Glu Phe Gly Xaa Gly Arg Ala Val Gln Gln Val Leu
 1 5 10 15

60

311

Asn Ile Glu Cys Leu Arg Asp Phe Leu Thr Pro Pro Leu Leu Ser Val
 20 25 30
 5 Arg Phe Arg Tyr Val Gly Ala Pro Gln Ala Leu Thr Leu Lys Leu Pro
 35 40 45
 Val Thr Xaa Asn Lys Phe Phe Gln Pro Thr Glu Met Ala Ala Gln Asp
 50 55 60
 10 Phe Phe Gln Arg Trp Lys Gln Leu Ser Leu Pro Gln Gln Glu Ala Gln
 65 70 75 80
 Lys Ile Phe Lys Ala Asn His Pro Met Asp Ala Glu Val Thr Lys Ala
 85 90 95
 15 Lys Leu Leu Gly Phe Gly Ser Ala Leu Leu Asp Asn Val Asp Pro Asn
 100 105 110
 Pro Glu Asn Phe Val Gly Ala Gly Ile Ile Gln Thr Lys Ala Leu Gln
 115 120 125
 Val Gly Cys Leu Leu Arg Leu Glu Pro Asn Ala Gln Ala Gln Met Tyr
 130 135 140
 25 Arg Leu Thr Leu Arg Thr Ser Lys Glu Pro Val Ser Arg His Leu Cys
 145 150 155 160
 Glu Leu Leu Ala Gln Gln Phe Xaa
 165
 30

(2) INFORMATION FOR SEQ ID NO: 185:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 185:

40 Met Phe Tyr Val Leu Ser Val Ser Pro Leu Leu Xaa Phe Leu Ala Cys
 1 5 10 15
 Gly Leu Cys Leu Cys Val Asn Trp Lys Ile Ala Ile Ser Gln Leu Ser
 20 25 30
 45 Leu Ser Phe Lys Asn Glu Leu Glu Lys Pro Xaa
 35 40

50

(2) INFORMATION FOR SEQ ID NO: 186:

- 55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 59 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

60 Met Lys Leu Phe Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly

312

1 5 10 15
 His Ile Leu Ala Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu
 20 25 30
 5 Gly Pro Gly Trp Val Pro Ser Ala Leu Xaa Arg Leu His Pro Gly His
 35 40 45
 10 Leu Ser Gly Ser Val Leu Val Ser Ala Ala Xaa
 50 55

(2) INFORMATION FOR SEQ ID NO: 187:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 189 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

Met Asp Val Asn Ile Ala Pro Leu Arg Ala Trp Asp Asp Phe Phe Pro
 5 10 15
 25 Gly Ser Asp Arg Phe Ala Arg Pro Asp Phe Arg Asp Ile Ser Lys Trp
 20 25 30
 Asn Asn Arg Val Val Ser Asn Leu Leu Tyr Tyr Gln Thr Asn Tyr Leu
 35 40 45
 30 Val Val Ala Ala Met Met Ile Ser Ile Val Gly Phe Leu Ser Pro Phe
 50 55 60
 35 Asn Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe
 65 70 75 80
 Val Trp Ala Ala His Asn Lys Asp Val Leu Arg Arg Met Lys Lys Arg
 85 90 95
 40 Tyr Pro Thr Thr Phe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu
 100 105 110
 Ile Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe
 115 120 125
 45 Pro Leu Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu
 130 135 140
 Lys Asn Lys Leu Glu Asn Lys Met Glu Gly Ile Gly Leu Lys Arg Thr
 145 150 155 160
 Pro Met Gly Ile Val Leu Asp Ala Leu Glu Gln Gln Glu Gly Ile
 165 170 175
 55 Asn Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu Xaa
 180 185

60

(2) INFORMATION FOR SEQ ID NO: 188:

313

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 146 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

5 Met Phe Leu Thr Arg Ile Leu Cys Pro Thr Tyr Ile Ala Leu Thr Phe
 1 5 10 15
 10 Leu Val Tyr Ile Val Ala Leu Val Ser Gly Gln Leu Cys Met Glu Ile
 20 25 30
 15 Ala Arg Gly Asn Ile Phe Phe Leu Asn Glu Leu Val Thr Thr Phe Cys
 35 40 45
 Cys Ser Cys Leu Leu Leu Ser Val Pro Tyr Leu His Pro Gly Phe Phe
 50 55 60
 20 Tyr Ser Ser Leu Cys Lys Cys Cys Phe Val Leu Val Val Leu Ser Arg
 65 70 75 80
 Ile Gly Ser Val Asn Glu Thr Trp Ser Cys Asn Phe Ser Ile Cys Ser
 85 90 95
 25 Tyr Leu Ile Phe Gly Ser Pro Ile Phe Thr Ala Val Ile Pro Lys Arg
 100 105 110
 30 Cys Ala Leu Glu Asp Ile Gln Asn Asn Pro Ile Gly Cys Leu Leu Arg
 115 120 125
 Cys Thr Pro Ala Trp Glu Thr Glu Gly Asp Ser Ile Ser Lys Lys Ile
 130 135 140
 35 Lys Lys
 145

40 (ii) INFORMATION FOR SEQ ID NO: 189:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

45 Met Gly Ser Arg Ala Glu Leu Cys Thr Leu Leu Gly Gly Phe Ser Phe
 1 5 10 15
 50 Leu Leu Leu Leu Ile Pro Gly Glu Gly Ala Lys Gly Gly Ser Leu Arg
 20 25 30
 55 Glu Ser Gln Gly Val Cys Ser Lys Gln Thr Leu Val Val Pro Leu His
 35 40 45
 Tyr Asn Glu Ser Tyr Ser Gln Pro Val Tyr Lys Pro Tyr Leu Thr Leu
 50 55 60
 60 Cys Ala Gly Ser Ala Ser Ala Ala Leu Thr Gly Pro Cys Thr Ala Leu

314

65 70 75 80

Cys Gly Gly Arg

5

(2) INFORMATION FOR SEQ ID NO: 190:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 190:

15

Met Met Gly Val Leu Gln Leu Leu His Ile Phe Trp Ala Tyr Leu Ile
1 5 10 15

20

Leu Arg Met Ala His Lys Phe Ile Thr Gly Lys Leu Val Glu Asp Glu
20 25 30

Arg Ser Thr Gly Lys Lys Gln Arg Ala Gln Arg Gly Arg Arg Leu Gln
35 40 45

25

Leu Gly Glu Glu Gln Arg Ala Gly Pro Xaa
50 55

30

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

35

Met Arg Arg Leu Val His Asp Leu Leu Pro Pro Glu Val Cys Ser Leu
1 5 10 15

40

Leu Asn Pro Ala Ala Ile Tyr Ala Asn Asn Glu Ile Ser Leu Arg Asp
20 25 30

45

Val Glu Val Tyr Gly Phe Asp Tyr Asp Tyr Thr Leu Ala Gln Tyr Ala
35 40 45

Asp Ala Leu His Pro Glu Ile Phe Ser Thr Ala Arg Asp Ile Leu Ile
50 55 60

50

Glu His Tyr Lys Tyr Pro Glu Gly Ile Arg Lys Tyr Asp Tyr Asn Pro
65 70 75 80

Ser Phe Ala Ile Arg Gly Leu His Tyr Asp Ile Gln Lys Ser Leu Leu
85 90 95

55

Met Lys Ile Asp Ala Phe His Tyr Val Gln Leu Gly Thr Ala Tyr Arg
100 105 110

60

Gly Leu Gln Pro Val Pro Asp Glu Glu Val Ile Glu Leu Tyr Gly Gly
115 120 125

315

Thr Gln His Ile Pro Leu Tyr Gln Met Ser Gly Phe Tyr Gly Lys Gly
 130 135 140

5 Pro Ser Ile Lys Gln Phe Met Asp Ile Phe Ser Leu Pro Glu Met Ala
 145 150 155 160

Leu Leu Ser Cys Val Val Asp Tyr Phe Leu Gly His Ser Leu Glu Phe
 165 170 175

10 Asp Gln Ala His Leu Tyr Lys Asp Val Thr Asp Ala Ile Arg Asp Val
 180 185 190

15 His Val Lys Gly Leu Met Tyr Gln Trp Ile Glu Gln Asp Met Glu Lys
 195 200 205

Tyr Ile Leu Arg Gly Asp Glu Thr Phe Ala Val Leu Ser Arg Leu Val
 210 215 220

20 Ala His Gly Lys Gln Leu Phe Leu Ile Thr Asn Ser Pro Phe Ser Phe
 225 230 235 240

Val Asp Lys Gly Met Arg His Met Val Gly Pro Asp Trp Arg His Ser
 245 250 255

25 Ser Met Trp Ser Leu Ser Arg Gln Thr Ser Pro Ala Ser Ser Leu Thr
 260 265 270

30 Gly Ala Ser Phe Xaa Glu Asn Ser Met Arg Arg Ala His Phe Ser Gly
 275 280 285

Thr Gly Ser Pro Ala Trp Lys Arg Ala Arg Ser Ile Gly Arg Glu Thr
 290 295 300

35 Cys Leu Thr Ser Tyr Ala Xaa
 305 310

40 (2) INFORMATION FOR SEQ ID NO: 192:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 318 amino acids

(B) TYPE: amino acid

45 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

Met Asn Trp Glu Leu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu
 1 5 10 15

50 Leu Leu Leu Val Gln Leu Leu Arg Phe Leu Arg Ala Asp Gly Asp Leu
 20 25 30

Thr Leu Leu Trp Ala Glu Trp Gln Gly Arg Arg Pro Glu Trp Glu Leu
 35 40 45

Thr Asp Met Val Val Trp Val Thr Gly Ala Ser Ser Gly Ile Gly Glu
 50 55 60

60 Glu Leu Ala Tyr Gln Leu Ser Lys Leu Gly Val Ser Leu Val Leu Ser

316

	65		70		75		80
	Ala Arg Arg Val His	Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Glu					
	85		90				95
5	Asn Gly Asn Leu Lys	Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu					
	100		105				110
10	Thr Asp Thr Gly Ser His	Glu Ala Ala Thr Lys Ala Val Leu Gln Glu					
	115		120				125
	Phe Gly Arg Ile Asp Ile	Leu Val Asn Asn Gly Gly Met Ser Gln Arg					
	130		135				140
15	Ser Leu Cys Met Asp Thr	Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu					
	145		150				155
	Leu Asn Tyr Leu Gly Thr	Val Ser Leu Thr Lys Cys Val Leu Pro His					
	165		170				175
20	Met Ile Glu Arg Lys Gln	Gly Lys Ile Val Thr Val Asn Ser Ile Leu					
	180		185				190
	Gly Ile Ile Ser Val Pro	Leu Ser Ile Gly Tyr Cys Ala Ser Lys His					
	195		200				205
25	Ala Leu Arg Gly Phe Phe	Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr					
	210		215				220
30	Pro Gly Ile Ile Val Ser	Asn Ile Cys Pro Gly Pro Val Gln Ser Asn					
	225		230				235
	Ile Val Glu Asn Ser Leu	Ala Gly Glu Val Thr Lys Thr Ile Gly Asn					
	245		250				255
35	Asn Gly Asp Gln Ser His	Lys Met Thr Thr Ser Arg Cys Val Arg Leu					
	260		265				270
40	Met Leu Ile Ser Met Ala	Asn Asp Leu Lys Glu Val Trp Ile Ser Glu					
	275		280				285
	Gln Pro Phe Leu Phe Ser	Asn Ile Phe Val Ala Ile His Ala Asn Leu					
	290		295				300
45	Gly Leu Val Asp Asn Gln	Gln Asp Gly Glu Glu Lys Asp Xaa					
	305		310				315

50 (2) INFORMATION FOR SEQ ID NO: 193:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:

Met Trp Pro Ser Phe Pro Gln Val Arg Val Gly Ser Phe Leu Phe Gly
1 5 10 15

60

317

Ile Leu Phe Phe Ser Phe Gly Ser Ser Ser Leu Pro Pro Gly Leu Pro
 20 25 30

Pro Pro Ala Ser Leu Leu Cys Cys Ala Val Gln Trp Gly Ala Arg Ala
 5 35 40 45

Leu Phe Leu Pro Ala
 50

10

(2) INFORMATION FOR SEQ ID NO: 194:

- (i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 42 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

20 Met Leu Val Thr Cys Ser Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala
 1 5 10 15

Ile Leu Ala Gln Leu Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu
 20 25 30

25 Thr Ile Trp Tyr Leu Lys Tyr His Trp Pro
 35 40

30

(2) INFORMATION FOR SEQ ID NO: 195:

- (i) SEQUENCE CHARACTERISTICS:
 35 (A) LENGTH: 102 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

40 Met Glu Gly Thr Glu Met Gly Ala Arg Pro Gly Gly His Pro Gln Lys
 1 5 10 15

Trp Ser Phe Leu Trp Ser Leu Ala Leu Trp Leu Pro Leu Ala Leu Ser
 20 25 30

45 Val Ser Leu Phe Leu Gly Leu Ser Leu Ser Pro Pro Gln Pro Gly Leu
 35 40 45

Ser Leu Trp Cys Thr Leu Ser Tyr Cys Cys Glu Gln Trp Lys Phe Lys
 50 55 60

50 Gly Thr Pro Ser Pro Ala Leu Leu Asn Leu Gly Thr Gln Pro Lys Lys
 65 70 75 80

55 Asp Lys Lys Leu Glu Asp Ser Ile Ala Thr Gln Leu Arg Glu Leu Pro
 85 90 95

Glu Lys Asn Ser Asn Xaa
 100

60

(2) INFORMATION FOR SEQ ID NO: 196:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 45 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

10 Met Ala Leu Thr Phe Leu Leu Val Leu Leu Thr Leu Ala Thr Ser Ala
1 5 10 15
His Gly Cys Thr Glu Thr Ser Asp Ala Gly Arg Ala Ser Thr Gly Gly
20 25 30
15 Pro Gln Arg Thr Ala Arg Thr Gln Trp Leu Leu Cys Xaa
35 40 45

(2) INFORMATION FOR SEQ ID NO: 197:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 355 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:

30 Met Gly Pro Ser Thr Pro Leu Leu Ile Leu Phe Leu Leu Ser Trp Ser
5 10 15
Gly Pro Leu Gln Gly Gln Gln His His Leu Val Glu Tyr Met Glu Arg
20 25 30
35 Arg Leu Ala Ala Leu Glu Glu Arg Leu Ala Gln Cys Gln Asp Gln Ser
35 40 45
Ser Arg His Ala Ala Glu Leu Arg Asp Phe Lys Asn Lys Met Leu Pro
50 55 60
40 Leu Leu Glu Val Ala Glu Lys Glu Arg Glu Ala Leu Arg Thr Glu Ala
65 70 75 80
45 Asp Thr Ile Ser Gly Arg Val Asp Arg Leu Glu Arg Glu Val Asp Tyr
85 90 95
Leu Glu Thr Gln Asn Pro Ala Leu Pro Cys Val Glu Phe Asp Glu Lys
100 105 110
50 Val Thr Gly Gly Pro Gly Thr Lys Gly Lys Gly Arg Arg Asn Glu Lys
115 120 125
Tyr Asp Met Val Thr Asp Cys Gly Tyr Thr Ile Ser Gln Val Arg Ser
130 135 140
55 Met Lys Ile Leu Lys Arg Phe Gly Gly Pro Ala Gly Leu Trp Thr Lys
145 150 155 160
60 Asp Pro Leu Gly Gln Thr Glu Lys Ile Tyr Val Leu Asp Gly Thr Gln
165 170 175

319

Asn Asp Thr Ala Phe Val Phe Pro Arg Leu Arg Asp Phe Thr Leu Ala
 180 185 190

5 Met Ala Ala Arg Lys Ala Ser Arg Val Arg Val Pro Phe Pro Trp Val
 195 200 205

Gly Thr Gly Gln Leu Val Tyr Gly Gly Phe Leu Tyr Phe Ala Arg Arg
 210 215 220

10 Pro Pro Gly Arg Pro Gly Gly Gly Gly Glu Met Glu Asn Thr Leu Gln
 225 230 235 240

Leu Ile Lys Phe His Leu Ala Asn Arg Thr Val Val Asp Ser Ser Val
 245 250 255

Phe Pro Ala Glu Gly Leu Ile Pro Pro Tyr Gly Leu Thr Ala Asp Thr
 260 265 270

20 Tyr Ile Asp Leu Ala Ala Asp Glu Gly Leu Trp Ala Val Tyr Ala
 275 280 285

Thr Arg Glu Asp Asp Arg His Leu Cys Leu Ala Lys Leu Asp Pro Gln
 290 295 300

25 Thr Leu Asp Thr Glu Gln Gln Trp Asp Thr Pro Cys Pro Arg Glu Asn
 305 310 315 320

Ala Glu Ala Ala Phe Val Ile Cys Gly Thr Leu Tyr Val Val Tyr Asn
 325 330 335

30 Thr Arg Pro Ala Ser Arg Ala Arg Ile Gln Cys Ser Phe Asp Ala Ser
 340 345 350

35 Gly Pro Xaa
 355

40 (2) INFORMATION FOR SEQ ID NO: 198:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 74 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198

Met Val Leu Pro Leu Leu Ile Phe Val Leu Leu Pro Lys Val Val Asn
 1 5 10 15

50 Thr Ser Asp Pro Asp Met Arg Arg Glu Met Glu Gln Ser Met Asn Met
 20 25 30

Leu Asn Ser Asn His Glu Leu Pro Asp Val Ser Glu Phe Met Thr Arg
 35 40 45

Leu Phe Ser Ser Lys Ser Ser Gly Lys Ser Ser Ser Gly Ser Ser Lys
 50 55 60

60 Thr Gly Lys Ser Gly Ala Gly Lys Arg Arg

320:

€€

70

5 (2) INFORMATION FOR SEQ ID NO: 199:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 113 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

Met Phe Thr Met Leu Cys Ile Asn Gly Thr Thr Pro Arg Pro Leu Pro
1 4 10 15

Val Pro Ser Pro Phe Gly Cys Met Ile Phe Phe Phe Phe Lys Asn Pro
20 25 30

Trp Lys Gln Arg Leu Leu Gln Gly Trp Leu Gly Ala Arg Pro Ile His
3' 40 41

Leu Leu Gly Tyr Leu Pro Leu Ser Leu Leu Trp Cys Pro Phe Pro Leu
50 55 60

Pro Cys Ala Arg Cys Ser Val Val Tyr Ile Ser Ser Pro Arg His Gly
65 70 75 80

Ala His Ala Pro Arg Asp Met Ile Leu Ser Leu Val Leu Ala His Gly
85 90 95

Ala Leu Tyr Lys Glu Leu Gly Gly Arg Gly Arg Lys Trp Glu Pro Ser
100 105 110

Xaā

(2) INFORMATION FOR SEQ ID NO: 200:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

Met Ala Cys Arg Cys Leu Ser Phe Leu Leu Met Gly Thr Phe Leu Ser
1 10 15

Val Ser Gln Thr Val Leu Ala Gln Leu Asp Ala Leu Leu Val Phe Pro
20 25 30

Gly Gln Val Ala Gln Leu Ser Cys Thr Leu Ser Pro Gln His Val Thr
35 40 45

Ile Arg Asp Tyr Gly Val Ser Trp Tyr Gln Gln Arg Ala Gly Ser Ala
50 55 60

Pro Arg Tyr Leu Leu Tyr Tyr Arg Ser Glu Glu Asp His His Arg Pro
65 70 75 80

321

Ala Asp Ile Pro Asp Arg Phe Ser Ala Ala Lys Asp Glu Ala His Asn
 85 90 95

5 Ala Cys Val Leu Thr Ile Ser Pro Val Gln Pro Glu Asp Asp Ala Asp
 100 105 110

Tyr Tyr Cys Ser Val Gly Tyr Gly Phe Ser Pro
 115 120

10

(2) INFORMATION FOR SEQ ID NO: 201:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 315 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

20 Met Ala Gly Gly Arg Cys Gly Pro Xaa Leu Thr Ala Leu Leu Ala Ala
 1 5 10 15

25 Trp Ile Ala Ala Val Ala Ala Thr Ala Gly Pro Glu Glu Ala Ala Leu
 20 25 30

Pro Pro Glu Gln Ser Arg Val Gln Pro Met Thr Ala Ser Asn Trp Thr
 35 40 45

30 Leu Val Met Glu Gly Glu Trp Met Leu Lys Phe Tyr Ala Pro Trp Cys
 50 55 60

Pro Ser Cys Gln Gln Thr Asp Ser Glu Trp Glu Ala Phe Ala Lys Asn
 65 70 75 80

35 Gly Glu Ile Leu Gln Ile Ser Val Gly Lys Val Asp Val Ile Gln Glu
 85 90 95

Pro Gly Leu Ser Gly Arg Phe Phe Val Thr Thr Leu Pro Ala Phe Phe
 100 105 110

His Ala Lys Asp Gly Ile Phe Arg Arg Tyr Arg Gly Pro Gly Ile Phe
 115 120 125

45 Glu Asp Leu Gln Asn Tyr Ile Leu Glu Lys Lys Trp Gln Ser Val Glu
 130 135 140

Pro Leu Thr Gly Trp Lys Ser Pro Ala Ser Leu Thr Met Ser Gly Met
 145 150 155 160

50 Ala Gly Leu Phe Ser Ile Ser Gly Lys Ile Trp His Leu His Asn Tyr
 165 170 175

Phe Thr Val Thr Leu Gly Ile Pro Ala Trp Cys Ser Tyr Val Phe Phe
 180 185 190

Val Ile Ala Thr Leu Val Phe Gly Leu Phe Met Gly Leu Val Leu Val
 195 200 205

60 Val Ile Ser Glu Cys Phe Tyr Val Pro Leu Pro Arg His Leu Ser Glu

323

	210		215		220
	Arg Ser Glu Gln Asn	Arg Arg Ser Glu Glu Ala His Arg Ala Glu Gln			
5	225	230	235	240	
	Leu Gln Asp Ala Glu Glu Glu Lys Asp Asp Ser Asn Glu Glu Glu Asn				
	245	250	255		
10	Lys Asp Ser Leu Val Asp Asp Glu Glu Glu Lys Glu Asp Leu Gly Asp				
	260	265	270		
	Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Asp Asn Leu Ala Ala Gly				
	275	280	285		
15	Val Asp Glu Glu Arg Ser Glu Ala Asn Asp Gln Gly Pro Pro Gly Glu				
	290	295	300		
	Asp Gly Val Thr Arg Glu Xaa Ser Arg Ala Xaa				
20	305	310	315		

(2) INFORMATION FOR SEQ ID NO: 202:

```

25      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 236 amino acids
          (B) TYPE: amino acid
          (D) TOPOLOGY: linear
      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

30      Met Gly Thr Ala Asp Ser Asp Glu Met Ala Pro Glu Ala Pro Gln His
          1           5           10           15

35      Thr His Ile Asp Val His Ile His Gln Glu Ser Ala Leu Ala Lys Leu
          20           25           30

          Leu Leu Thr Cys Cys Ser Ala Leu Arg Pro Arg Ala Thr Gln Ala Arg
          35           40           45

40      Gly Ser Ser Arg Leu Leu Val Ala Ser Trp Val Met Gln Ile Val Leu
          50           55           60

          Gly Ile Leu Ser Ala Val Leu Gly Gly Phe Phe Tyr Ile Arg Asp Tyr
          65           70           75           80

45      Thr Leu Leu Val Thr Ser Gly Ala Ala Ile Trp Thr Gly Ala Val Ala
          85           90           95

          Val Leu Ala Gly Ala Ala Ala Phe Ile Tyr Glu Lys Arg Gly Gly Thr
          100          105          110

50      Tyr Trp Ala Leu Leu Arg Thr Leu Leu Ala Leu Ala Ala Phe Ser Thr
          115          120          125

          Ala Ile Ala Ala Leu Lys Leu Trp Asn Glu Asp Phe Arg Tyr Gly Tyr
          130          135          140

55      Ser Tyr Tyr Asn Ser Ala Cys Arg Ile Ser Ser Ser Ser Asp Trp Asn
          145          150          155          160

```

60

323

Thr Pro Ala Pro Thr Gln Ser Pro Glu Glu Val Arg Arg Leu His Leu
165 170 175

5 Cys Thr Ser Phe Met Asp Met Leu Lys Ala Leu Phe Arg Thr Leu Gln
180 185 190

Ala Met Leu Leu Gly Val Trp Ile Leu Leu Leu Leu Ala Ser Leu Ala
195 200 205

10 Pro Leu Trp Leu Tyr Cys Trp Arg Met Phe Pro Thr Lys Gly Lys Arg
210 215 220

Asp Gln Lys Glu Met Leu Glu Val Ser Gly Ile Xaa
225 230 235

15

(2) INFORMATION FOR SEQ ID NO: 203:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

25 Met Ile His Leu Gly His Ile Leu Phe Leu Leu Leu Leu Pro Val Ala
1 5 10 15

30 Ala Ala Gln Thr Thr Pro Gly Glu Arg Ser Ser Leu Pro Ala Phe Tyr
20 25 30

Pro Gly Thr Ser Gly Ser Cys Ser Gly Cys Gly Ser Leu Ser Leu Pro
35 40 45

35 Leu Leu Ala Gly Leu Val Ala Ala Asp Ala Val Ala Ser Leu Leu Ile
50 55 60

Val Gly Ala Val Phe Leu Cys Ala Arg Pro Arg Arg Ser Pro Ala Gln
65 70 75 80

40 Glu Asp Gly Lys Val Tyr Ile Asn Met Pro Gly Arg Gly
85 90

45

(2) INFORMATION FOR SEQ ID NO: 204:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 204:

55 Met Trp Ser Ala Gly Arg Gly Gly Ala Ala Trp Pro Val Leu Leu Gly
1 5 10 15

Leu Leu Leu Ala Leu Leu Val Pro Gly Gly Gly Ala Ala Lys Thr Gly
20 25 30

60 Ala Asp Ser

35

5 (2) INFORMATION FOR SEQ ID NO: 205:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 amino acids

(B) TYPE: amino acid

10 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

Asp Cys Xaa His Val Ser Val Leu Gln Ser Thr Ile Ser Pro Leu Leu
 1 5 10 15
 Pro Leu Pro Leu Leu Leu Pro His Gly Asn Cys Glu Glu Ala Pro Trp
 20 25 30
 Gln Ala Ala Val Ile Gly Gly Gly Asp Arg Ile
 35 40

25 (2) INFORMATION FOR SEQ ID NO: 206:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 85 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

Met Arg Asp Cys Leu Ser Leu Lys Pro Arg Pro Leu Phe Pro Thr Gln
 1 5 10 15
 Phe Phe Phe Ile Leu Leu Leu Ile Phe Ile Ala Glu Val Ala Ala Ala
 20 25 30
 Val Val Ala Leu Val Tyr Thr Thr Met Val Arg His Trp Asp Gly Gly
 35 40 45
 Arg Glu Glu Asp Trp Ala Lys Pro Trp Glu Trp Ala Val Ala Cys Glu
 50 55 60
 Trp Pro Pro Ser Val Pro Ala Pro Lys His Trp Pro Ala Ser Pro Arg
 65 70 75 80
 Leu Ser Thr Ser Xaa
 85

50

(2) INFORMATION FOR SEQ ID NO: 207:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 208 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

60 Met His Gly Asn Glu Ala Leu Gly Arg Glu Leu Leu Leu Leu Met

326

	1	5	10	15
	Gln Phe Leu Cys His Glu Phe Leu Arg Xaa Asn Pro Arg Val Thr Arg	20	25	30
5	Leu Leu Ser Glu Met Arg Ile His Leu Leu Pro Ser Met Asn Pro Asp	35	40	45
10	Gly Tyr Glu Ile Ala Tyr His Arg Gly Ser Glu Leu Val Gly Trp Ala	50	55	60
	Glu Gly Arg Trp Asn Asn Gln Ser Ile Asp Leu Asn His Asn Phe Ala	65	70	75
15	Xaa Leu Asn Thr Pro Leu Trp Glu Ala Gln Asp Asp Gly Lys Val Pro	85	90	95
	His Ile Val Pro Asn His His Leu Pro Leu Pro Thr Tyr Tyr Thr Leu	100	105	110
20	Pro Asn Ala Thr Val Ala Pro Glu Thr Arg Ala Val Ile Lys Trp Met	115	120	125
	Lys Arg Ile Pro Phe Val Leu Ser Ala Asn Leu His Gly Gly Glu Leu	130	135	140
25	Val Val Ser Tyr Pro Phe Asp Met Thr Arg Thr Pro Trp Ala Ala Arg	145	150	155
30	Glu Leu Thr Pro Thr Pro Asp Asp Ala Val Phe Arg Trp Leu Ser Thr	165	170	175
	Val Tyr Ala Gly Ser Asn Leu Ala Met Gln Asp Thr Ser Arg Arg Pro	180	185	190
35	Cys His Ser Gln Asp Phe Ser Val His Gly Asn Ile Ile Asn Gly Ala	195	200	205

40

(2) INFORMATION FOR SEQ ID NO: 208:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(E) TYPE: amino acid

(E) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208

Met Glu Ile Ser Cys Leu Leu Leu Leu Ile Gln Asp Ser Asp Glu Met
1 5 10 15

55 Glu Asp Gly Pro Gly Val Gln Asp
 20

60 (2) INFORMATION FOR SEQ ID NO: 209:

326

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 483 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

5 Met Ala Thr Gly Gly Gly Ile Arg Ala Met Thr Ser Leu Tyr Gly Gln
 1 5 1 15
 10 Leu Ala Gly Leu Lys Glu Leu Gly Leu Leu Asp Cys Xaa Ser Tyr Ile
 20 25 30
 15 Thr Gly Ala Ser Gly Ser Thr Trp Ala Leu Ala Asn Leu Tyr Lys Asp
 35 40 45
 Pro Glu Trp Ser Gln Lys Asp Leu Ala Gly Pro Thr Glu Leu Leu Lys
 50 55 60
 20 Thr Gln Val Thr Lys Asn Lys Leu Gly Val Leu Ala Pro Ser Gln Leu
 65 70 75 80
 Gln Arg Tyr Arg Gln Glu Leu Ala Glu Arg Ala Arg Leu Gly Tyr Pro
 85 90 95
 25 Ser Cys Phe Thr Asn Leu Trp Ala Leu Ile Asn Glu Ala Leu Leu His
 100 105 110
 Asp Glu Pro His Asp His Lys Leu Ser Asp Gln Arg Glu Ala Leu Ser
 115 120 125
 30 His Gly Gln Asn Pro Leu Pro Ile Tyr Cys Ala Leu Asn Thr Lys Gly
 130 135 140
 35 Gln Ser Leu Thr Thr Phe Glu Phe Gly Glu Trp Cys Glu Phe Ser Pro
 145 150 155 160
 Tyr Glu Val Gly Phe Pro Lys Tyr Gly Ala Phe Ile Pro Ser Glu Leu
 165 170 175
 40 Phe Gly Ser Glu Phe Phe Met Gly Gln Leu Met Lys Arg Leu Pro Glu
 180 185 190
 Ser Arg Ile Cys Phe Leu Glu Gly Ile Trp Ser Asn Leu Tyr Ala Ala
 195 200 205
 45 Asn Leu Gln Asp Ser Leu Tyr Trp Ala Ser Glu Pro Ser Gln Phe Trp
 210 215 220
 50 Asp Arg Trp Val Arg Asn Gln Ala Asn Leu Asp Lys Glu Gln Val Pro
 225 230 235 240
 Leu Leu Lys Ile Glu Glu Pro Pro Ser Thr Ala Gly Arg Ile Ala Glu
 245 250 255
 55 Phe Phe Thr Asp Leu Leu Thr Trp Arg Pro Leu Ala Gln Ala Thr His
 260 265 270
 60 Asn Phe Leu Arg Gly Leu His Phe His Lys Asp Tyr Phe Gln His Pro
 275 280 285

327

His Phe Ser Thr Trp Lys Ala Thr Thr Leu Asp Gly Leu Pro Asn Gln
 290 295 300
 5 Leu Thr Pro Ser Glu Pro His Leu Cys Leu Leu Asp Val Gly Tyr Leu
 305 310 315 320
 Ile Asn Thr Ser Cys Leu Pro Leu Leu Gln Pro Thr Arg Asp Val Asn
 325 330 335
 10 Leu Ile Leu Ser Leu Asp Tyr Asn Leu His Gly Ala Phe Gln Gln Leu
 340 345 350
 Gln Leu Leu Gly Arg Phe Cys Gln Glu Gln Gly Ile Pro Phe Pro Pro
 355 360 365
 15 Ile Ser Pro Ser Pro Glu Glu Gln Leu Gln Pro Arg Glu Cys His Thr
 370 375 380
 20 Phe Ser Asp Pro Thr Cys Pro Gly Ala Pro Ala Val Leu His Phe Pro
 385 390 395 400
 Leu Val Ser Asp Ser Phe Arg Glu Tyr Ser Ala Pro Gly Val Arg Arg
 405 410 415
 25 Thr Pro Glu Glu Ala Ala Ala Gly Glu Val Asn Leu Ser Ser Ser Asp
 420 425 430
 Ser Pro Tyr His Tyr Thr Lys Val Thr Tyr Ser Gln Glu Asp Val Asp
 435 440 445
 30 Lys Leu Leu His Leu Thr His Tyr Asn Val Cys Asn Asn Gln Glu Gln
 450 455 460
 35 Leu Leu Glu Ala Leu Arg Gln Ala Val Gln Arg Arg Arg Gln Arg Arg
 465 470 475 480
 Pro His Xaa

40

(2) INFORMATION FOR SEQ ID NO: 210:

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

50

Leu Glu Val Gly Cys Ile Gln Val Ala Pro Asp Thr Phe

1

5

10

55

(2) INFORMATION FOR SEQ ID NO: 211:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids

60

(B) TYPE: amino acid

328

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

Met Ser Leu Phe Phe Leu Leu Thr Leu Ile Ser Lys Leu His Gly Asp
 5 1 5 10 15
 Ala Glu Val Cys
 20

10

(2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 55 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

20 Met Pro His Pro Pro Leu Pro Glu Thr Ser Leu Glu Ala Gln Leu Pro
 1 5 10 15
 Met Gly Leu Leu Gln Leu Leu Arg Cys Ser Val Gln Ala Trp Ser Pro
 20 25 30
 25 Pro Pro Ser Ser Phe Cys Pro Gly Ser Glu Pro Arg Ser Ala Ser Ala
 35 40 45
 His Trp Gly Tyr Trp Trp Pro
 30 50 55

35

(2) INFORMATION FOR SEQ ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:

 (A) LENGTH: 35 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:

40 Asp Pro Glu Thr Arg Trp His His Gly Gly Ser Ala Gln Asn Gly Leu
 1 5 10 15
 45 Leu Met Leu Ile Ser Val Leu Gln Gln Pro Val Ile Gly Thr Gly Ser
 20 25 30
 Tyr Leu Cys
 35

50

(2) INFORMATION FOR SEQ ID NO: 214:

(i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 230 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

60

329

Met Glu Pro Leu Arg Leu Leu Ile Leu Leu Phe Val Thr Glu Leu Ser
 1 5 10 15

Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu
 5 20 25 30

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys
 35 40 45

Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val
 10 50 55 60

Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly
 65 70 75 80

Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr
 15 85 90 95

Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser
 20 100 105 110

Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val
 115 120 125

Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro
 25 130 135 140

Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser
 145 150 155 160

Arg Ser Leu Leu Glu Gly Glu Ile Pro Phe Pro Pro Thr Ser Ile Leu
 30 165 170 175

Leu Leu Leu Ala Cys Ile Phe Leu Ile Lys Ile Leu Ala Ala Ser Xaa
 35 180 185 190

Leu Trp Ala Ala Ala Trp His Gly Gln Lys Pro Gly Thr His Pro Pro
 195 200 205

Ser Glu Leu Asp Cys Gly His Asp Pro Gly Tyr Gln Leu Gln Thr Leu
 40 210 215 220

Pro Gly Leu Arg Asp Thr
 225 230

45

(2) INFORMATION FOR SEQ ID NO: 215:

50

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 231 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

60

Met Glu Pro Leu Arg Leu Leu Ile Leu Leu Phe Val Thr Glu Leu Ser
 1 5 10 15

Gly Ala His Asn Thr Thr Val Phe Gln Gly Val Ala Gly Gln Ser Leu
 20 25 30

330

Gln Val Ser Cys Pro Tyr Asp Ser Met Lys His Trp Gly Arg Arg Lys
 35 40 45
 5 Ala Trp Cys Arg Gln Leu Gly Glu Lys Gly Pro Cys Gln Arg Val Val
 50 55 60
 Ser Thr His Asn Leu Trp Leu Leu Ser Phe Leu Arg Arg Trp Asn Gly
 65 70 75 80
 10 Ser Thr Ala Ile Thr Asp Asp Thr Leu Gly Gly Thr Leu Thr Ile Thr
 85 90 95
 Leu Arg Asn Leu Gln Pro His Asp Ala Gly Leu Tyr Gln Cys Gln Ser
 100 105 110
 Leu His Gly Ser Glu Ala Asp Thr Leu Arg Lys Val Leu Val Glu Val
 115 120 125
 20 Leu Ala Asp Pro Leu Asp His Arg Asp Ala Gly Asp Leu Trp Phe Pro
 130 135 140
 Gly Glu Ser Glu Ser Phe Glu Asp Ala His Val Glu His Ser Ile Ser
 145 150 155 160
 25 Arg Ser Leu Leu Glu Gly Glu Ile Pro Phe Pro Pro Thr Ser Ile Leu
 165 170 175
 Leu Leu Leu Ala Cys Ile Phe Leu Ile Lys Ile Leu Ala Ala Ser Ala
 180 185 190
 Leu Trp Ala Ala Ala Trp His Gly Gln Lys Pro Gly Thr His Pro Pro
 195 200 205
 35 Ser Glu Leu Asp Cys Gly His Asp Pro Gly Tyr Gln Leu Gln Thr Leu
 210 215 220
 Pro Gly Leu Arg Asp Thr Xaa
 225 230
 40

(2) INFORMATION FOR SEQ ID NO: 216:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 127 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:
 50 Met Gly Leu Thr Gly Phe Gly Val Phe Phe Leu Phe Phe Gly Met Ile
 5 10 15
 Leu Phe Phe Asp Lys Ala Leu Leu Ala Ile Gly Asn Val Leu Phe Val
 20 25 30
 Ala Gly Leu Ala Phe Val Ile Gly Leu Glu Arg Thr Phe Arg Phe Phe
 35 40 45
 60 Phe Gln Lys His Lys Met Lys Ala Thr Gly Phe Phe Leu Gly Gly Val

331

[illegible]

15

(2) INFORMATION FOR SEQ ID NO: 217:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

25 Met Ile Arg Lys Leu His Lys Ile Ile Val Phe Ser Pro Arg Val Ile
1 5 10 15

Val Leu Leu Asn Cys Phe Phe Phe Ile Lys Ala Lys Phe Val Leu Tyr
20 25 30

30 Ile Phe Val Phe His Val Leu Asp Gly Ser Ile Ser Tyr Pro Val
35 40 45

35

(2) INFORMATION FOR SEQ ID NO: 218:

40 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218

45 Met Leu Leu Asn Gln His Phe Lys Ile Phe Gly Ser Leu Ile His Met
1 10 15
Asn Leu Leu Phe Ala Leu Ile Ser Leu Gly Ser Ser Asn Leu Ser Gly
20 25 30

50 Val Gln Phe Cys Cys Glu Thr Val Gln.
35 40

55 (2) INFORMATION FOR SEQ ID NO: 219:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

330

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

Met Gln Pro Leu Asn Phe Ser Ser Thr Xaa Cys Ser Ser Phe Ser Pro
 1 5 10 15
 5 Pro Thr Thr Val Ile Leu Leu Ile Leu Leu Cys Phe Glu Gly Leu Leu
 20 25 30
 10 Phe Leu Ile Phe Thr Ser Val Met Phe Gly Thr Gln Val His Ser Ile
 35 40 45
 Cys Thr Asp Glu Thr Gly Ile Glu Gln Leu Lys Lys Glu Glu Arg Arg
 50 55 60
 15 Trp Ala Lys Lys Thr Lys Trp Met Asn Met Lys Ala Val Phe Gly His
 65 70 75 80
 Pro Phe Ser Leu Gly Trp Ala Ser Pro Phe Ala Thr Pro Asp Gln Gly
 85 90 95
 20 Lys Ala Asp Pro Tyr Gln Tyr Val Val
 100 105

25

(2) INFORMATION FOR SEQ ID NO: 220:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 29 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

Met Tyr Thr Asn His Phe Asn Leu Tyr Leu Lys Tyr Ile Leu Leu Ile
 1 5 10 15
 35 Ile Leu Ile Leu Asn Met Thr Asn Ser Ser Ser Arg Tyr
 20 25

40

(2) INFORMATION FOR SEQ ID NO: 221:

(i) SEQUENCE CHARACTERISTICS:

45

- (A) LENGTH: 17 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

Met Asn Glu Leu Leu Leu Phe Phe Phe Phe Phe Phe Phe Leu His Phe
 1 5 10 15
 Val

55

(2) INFORMATION FOR SEQ ID NO: 222:

60

(i) SEQUENCE CHARACTERISTICS:

335

(A) LENGTH: 138 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

5

Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu
 1 10 15

10

Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala
 20 25 30

Gln Glu Ala Gly Thr Ser Lys Pro Asn Glu Glu Ile Ser Gly Pro Ala
 35 40 45

15

Glu Pro Ala Ser Pro Pro Glu Thr Thr Thr Thr Ala Gln Glu Xaa Ser
 50 55 60

Ala Ala Ala Val Gln Gly Thr Ala Lys Val Thr Ser Ser Arg Gln Glu
 65 70 75 80

20

Leu Asn Pro Leu Lys Ser Ile Val Glu Lys Ser Ile Leu Leu Thr Glu
 85 90 95

25

Gln Ala Leu Ala Lys Ala Gly Lys Gly Met His Gly Gly Val Pro Gly
 100 105 110

Gly Lys Gln Phe Ile Glu Asn Gly Ser Glu Phe Ala Gln Lys Leu Leu
 115 120 125

30

Lys Lys Phe Ser Leu Leu Lys Pro Trp Ala
 130 135

35

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

Met Leu Gly Cys Gly Ile Pro Ala Leu Gly Leu Leu Leu Leu Gln
 1 5 10 15

45

Xaa Ser Ala Asp Gly Asn Gly Ile Gln Gly Phe Phe Tyr Pro Trp Ser
 20 25 30

50

Cys Glu Gly Asp Ile Trp Asp Arg Glu Ser Cys Gly Gly Gln Ala Ala
 35 40 45

Ile Arg
 50

55

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

60

33c

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

5 Met Glu Ala Val Phe Thr Val Phe Phe Phe Leu Leu Phe Cys Phe
 1 5 10 15

10 (2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 155 amino acids

(E) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

Met Gly Phe Gly Ala Thr Leu Ala Val Gly Leu Thr Ile Phe Val Leu
 1 5 10 15

20 Ser Val Val Thr Ile Ile Ile Cys Phe Thr Cys Ser Cys Cys Cys Leu
 20 25 30

25 Tyr Lys Thr Cys Arg Arg Pro Arg Pro Val Val Thr Thr Thr Thr Ser
 35 40 45

Thr Thr Val Val His Ala Pro Tyr Pro Gln Pro Pro Ser Val Pro Pro
 50 55 60

30 Ser Tyr Pro Gly Pro Ser Tyr Gln Gly Tyr His Thr Met Pro Pro Gln
 65 70 75 80

Pro Gly Met Pro Ala Ala Pro Tyr Pro Met Gln Tyr Pro Pro Pro Tyr
 85 90 95

35 Pro Ala Gln Pro Met Gly Pro Pro Ala Tyr His Glu Thr Leu Ala Gly
 100 105 110

40 Gly Ala Ala Ala Pro Tyr Pro Ala Ser Gln Pro Pro Tyr Asn Pro Xaa
 115 120 125

Tyr Met Asp Ala Pro Lys Xaa Xaa Ser Glu His Ser Leu Ala Ser Leu
 130 135 140

45 Ala Ala Thr Trp Leu Cys Cys Val Cys Ala Xaa
 145 150 155

50 (2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(E) TYPE: amino acid

55 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

60 Met Gly Phe Gly Ala Thr Leu Ala Val Gly
 1 5 10

(2) INFORMATION FOR SEQ ID NO: 227:

- 5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

10 Met Ser Ile Phe Leu Val Met Ser Ile Ser Cys Ser Ser Thr Ser His
 1 5 10 15
 15 Cys Tyr Ser Phe
 20

(2) INFORMATION FOR SEQ ID NO: 228:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 94 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
 1 5 10 15
 30 Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys
 20 25 30
 Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val
 35 40 45
 Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu
 50 55 60
 40 Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser
 65 70 75 80
 Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr Xaa
 85 90

(2) INFORMATION FOR SEQ ID NO: 229:

- 50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 94 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

55 Met Ser Phe Ser Phe Ile Ile Phe Leu Leu Leu Val Cys Gln Glu Ile
 1 5 10 15
 Thr Phe Cys Met Ser Tyr Gly Asp Ala Val Asn Cys Phe Ser Glu Cys
 20 25 30
 60

336

Phe Ser Asn Leu Gln Thr Ile Tyr Ile Ser Cys Leu Gln His Ala Val
 35 40 45
 Cys Lys His Ser Val Ile Trp Ser Ile Gln Leu Phe Val Arg Ala Leu
 5 50 55 60
 Pro Ile Ser Lys Cys Ala Glu Leu Ser Ile Asp Gly Ile Phe Arg Ser
 65 70 75 80
 10 Phe His Glu Asn Trp Lys Cys Ser Trp Val Ala Pro Thr Xaa
 85 90

15 (2) INFORMATION FOR SEQ ID NO: 230:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230.

Met Gly Trp Ser Ala Gly Leu Leu Phe Leu Leu Ile Leu Tyr Leu Pro
 1 5 10 15
 25 Val Pro Gly Trp Met Glu Arg Glu Asp Gly Gly Asp Gly Thr Ser Phe
 20 25 30
 Thr Ser Gly Ser Trp
 30 35

35 (2) INFORMATION FOR SEQ ID NO: 231:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 81 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231.

Met Ala Thr Leu Trp Gly Gly Leu Leu Arg Leu Gly Ser Leu Leu Ser
 1 5 10 15
 45 Leu Ser Cys Leu Ala Leu Ser Val Leu Leu Ala His Val Gln Thr
 20 25 30
 Pro Pro Arg Ile Ser Arg Met Ser Asp Val Asn Val Ser Ala Leu Pro
 35 40 45
 50 Ile Lys Lys Ile Leu Gly Ile Phe Ile Ile Arg Thr Tyr Leu Arg Lys
 50 55 60
 Ile Val Ile Ala Phe Met Leu Trp Ser Pro Cys Leu Cys Gly Gly Leu
 55 65 70 75 80
 Met

60

337

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 301 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

5 Met Asp Ala Arg Trp Trp Ala Val Val Val Leu Ala Ala Phe Pro Ser
 1 5 10 15
 Leu Gly Ala Gly Gly Glu Thr Pro Glu Ala Pro Pro Glu Ser Trp Thr
 20 25 30
 15 Gln Leu Trp Phe Phe Arg Phe Val Val Asn Ala Ala Gly Tyr Ala Xaa
 35 40 45
 Phe Met Val Pro Gly Tyr Leu Leu Val Gln Tyr Phe Arg Arg Lys Asn
 20 50 55 60
 Tyr Leu Glu Thr Gly Arg Gly Leu Cys Phe Pro Leu Val Lys Ala Cys
 65 70 75 80
 25 Val Phe Gly Asn Glu Pro Lys Ala Ser Asp Glu Val Pro Leu Ala Pro
 85 90 95
 Arg Thr Glu Ala Ala Glu Thr Thr Pro Met Trp Gln Ala Leu Lys Leu
 100 105 110
 30 Leu Phe Cys Ala Thr Gly Leu Gln Val Ser Tyr Leu Thr Trp Gly Val
 115 120 125
 Leu Gln Glu Arg Val Met Thr Arg Ser Tyr Gly Ala Thr Ala Thr Ser
 35 130 135 140
 Pro Gly Glu Arg Phe Thr Asp Ser Gln Phe Leu Val Leu Met Asn Arg
 145 150 155 160
 40 Val Leu Ala Leu Ile Val Ala Gly Leu Ser Cys Val Leu Cys Lys Gln
 165 170 175
 Pro Arg His Gly Ala Pro Met Tyr Arg Tyr Ser Phe Ala Ser Leu Ser
 180 185 190
 45 Asn Val Leu Ser Ser Trp Cys Gln Tyr Glu Ala Leu Lys Phe Val Ser
 195 200 205
 Phe Pro Thr Gln Val Leu Ala Lys Ala Ser Lys Val Ile Pro Val Met
 50 210 215 220
 Leu Met Gly Lys Leu Val Ser Arg Arg Xaa Asn Glu His Trp Glu Tyr
 225 230 235 240
 55 Leu Thr Ala Thr Leu Ile Ser Ile Gly Val Ser Met Phe Leu Leu Ser
 245 250 255
 Ser Gly Pro Glu Pro Arg Ser Ser Pro Ala Thr Thr Leu Ser Gly Leu
 260 265 270
 60

338

Ile Leu Leu Ala Gly Tyr Ile Ala Phe Asp Ser Phe Thr Ser Asn Trp
275 280 285

5 Gln Asp Ala Cys Leu Pro Ile Arg Cys His Arg Cys Arg
290 295 300

10 (2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 amino acid
(B) TYPE: amino acid
(D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

Met Ser Asp Leu Leu Leu Leu Gly Leu Ile Gly Gly Leu Thr Leu Leu
1 5 10 15

20 Leu Leu Leu Thr Leu Leu Ala Phe Ala Gly Tyr Ser Gly Leu Leu Ala
20 25 30

Gly Val Glu Val Ser Ala Gly Ser Pro Pro Ile Arg Asn Val Thr Val
35 40 45

25 Ala Tyr Lys Phe His Met Gly Leu Tyr Gly Glu Thr Gly Arg Leu Phe
50 55 60

30 Thr Glu Ser Cys Ser Ile Ser Pro Lys Leu Arg Ser Ile Ala Val Tyr
65 70 75 80

Tyr Asp Asn Pro His Met Val Pro Pro Asp Lys Cys Arg Cys Ala Val
85 90 95

35 Gly Ser Ile Leu Ser Glu Gly Glu Glu Ser Pro Ser Pro Glu Leu Ile
100 105 110

Asp Leu Tyr Gln Lys Phe Gly Phe Lys Val Phe Ser Phe Pro Ala Pro
115 120 125

40 Ser His Val Val Thr Ala Thr Phe Pro Tyr Thr Thr Ile Leu Ser Ile
130 135 140

45 Trp Leu Ala Thr Arg Arg Val His Pro Ala Leu Asp Thr Tyr Ile Lys
145 150 155 160

Glu Arg Lys Leu Cys Ala Tyr Pro Arg Leu Glu Ile Tyr Gln Glu Asp
165 170 175

50 Gln Ile His Phe Met Cys Pro Leu Ala Xaa Gln Gly Asp Phe Tyr Val
180 185 190

Pro Glu Met Lys Glu Thr Glu Trp Lys Trp Arg Gly Leu Val Glu Ala
195 200 205

55 Ile Asp Thr Gln Val Asp Gly Thr Gly Ala Asp Thr Met Ser Asp Thr
210 215 220

60 Ser Ser Val Ser Leu Glu Val Ser Pro Gly Ser Arg Glu Thr Ser Ala
225 230 235 240

336

Ala Thr Leu Ser Pro Gly Ala Ser Ser Arg Gly Trp Asp Asp Gly Asp
 245 250 255

5 Thr Arg Ser Glu His Ser Tyr Ser Glu Ser Gly Ala Ser Gly Ser Ser
 260 265 270

Phe Glu Glu Leu Asp Leu Glu Gly Glu Gly Pro Leu Gly Glu Ser Arg
 275 280 285

10 Leu Asp Pro Gly Thr Xaa Pro Leu Gly Thr Thr Lys Trp Leu Trp Glu
 290 295 300

Pro Thr Ala Pro Glu Lys Gly Lys Glu
 15 305 310

(2) INFORMATION FOR SEQ ID NO: 234:
 20

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

Pro Gln Ser Leu Ile Leu His Leu Leu Leu Phe Phe Phe Leu Leu Phe
 1 5 10 15

30 Leu Phe Phe Ile Phe Ile Phe Leu Phe Phe Leu Gln Cys Leu Thr Phe
 20 25 30

Leu Phe Xaa Lys Pro Arg Gly Arg Tyr His Gly Leu Cys Phe Lys Phe
 35 40 45

40 (2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

Pro Ala Leu Arg Pro Ala Leu Leu Trp Ala Leu Leu Ala Leu Trp Leu
 50 1 5 10 15

Cys Cys Ala Thr Pro Arg Met His Cys Ser Val Glu Met Ala Met Asp
 20 25 30

55 Pro Val

60 (2) INFORMATION FOR SEQ ID NO: 236:

346

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 313 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

5 Met Thr Arg Gly Gly Pro Gly Gly Arg Pro Gly Leu Pro Gln Pro Pro
 1 5 1 11
 10 Pro Leu Leu Leu Leu Leu Leu Leu Xaa Leu Leu Leu Val Thr Ala Glu
 20 25 30
 15 Pro Pro Lys Pro Ala Gly Val Tyr Tyr Ala Thr Ala Tyr Trp Met Pro
 35 40 45
 Ala Glu Lys Thr Val Gln Val Lys Asn Val Met Asp Lys Asn Gly Asp
 50 55 60
 20 Ala Tyr Gly Phe Tyr Asn Asn Ser Val Lys Thr Thr Gly Trp Gly Ile
 65 70 75 80
 Leu Glu Ile Arg Ala Gly Tyr Gly Ser Gln Thr Leu Ser Asn Glu Ile
 85 90 95
 25 Ile Met Phe Val Ala Gly Phe Leu Glu Gly Tyr Leu Thr Ala Pro His
 100 105 110
 30 Met Asn Asp His Tyr Thr Asn Leu Tyr Pro Gln Leu Ile Thr Lys Pro
 115 120 125
 Ser Ile Met Asp Lys Val Gln Asp Phe Met Glu Lys Gln Asp Lys Trp
 130 135 140
 35 Thr Arg Lys Asn Ile Lys Glu Tyr Lys Thr Asp Ser Phe Trp Arg His
 145 150 155 160
 Thr Gly Tyr Val Met Ala Gln Ile Asp Gly Leu Tyr Val Gly Ala Lys
 165 170 175
 40 Lys Arg Ala Ile Leu Glu Gly Thr Lys Pro Met Thr Leu Phe Gln Ile
 180 185 190
 45 Gln Phe Leu Asn Ser Val Gly Asp Leu Leu Asp Leu Ile Pro Ser Leu
 195 200 205
 Ser Pro Thr Lys Asn Gly Ser Leu Lys Val Phe Lys Arg Trp Asp Met
 210 215 220
 50 Gly His Cys Ser Ala Leu Ile Lys Val Leu Pro Gly Phe Glu Asn Ile
 225 230 235 240
 Leu Phe Ala His Ser Ser Trp Tyr Thr Tyr Ala Ala Met Leu Arg Ile
 245 250 255
 55 Tyr Lys His Trp Asp Phe Asn Xaa Ile Asp Lys Asp Thr Ser Ser Ser
 260 265 270
 60 Arg Leu Ser Phe Ser Ser Tyr Pro Gly Phe Leu Glu Ser Leu Asp Asp
 275 280 285

341

Phe Tyr Ile Leu Ser Ser Gly Leu Ile Leu Leu Gln Thr Thr Asn Ser
 290 295 300

5 Val Phe Asn Lys Thr Leu Leu Lys Gln
 305 310

10 (2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 296 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

Met Leu Gln Gly Pro Gly Ser Leu Leu Leu Leu Phe Leu Ala Ser His
 1 5 10 15

20 Cys Cys Leu Gly Ser Ala Arg Gly Leu Phe Leu Phe Gly Gln Pro Asp
 20 25 30

25 Phe Ser Tyr Lys Arg Xaa Asn Cys Lys Pro Ile Pro Val Asn Leu Gln
 35 40 45

Leu Cys His Gly Ile Glu Tyr Gln Asn Met Arg Leu Pro Asn Leu Leu
 50 55 60

30 Gly His Glu Thr Met Lys Glu Val Leu Glu Gln Ala Gly Ala Trp Ile
 65 70 75 80

Pro Leu Val Met Lys Gln Cys His Pro Asp Thr Lys Lys Phe Leu Cys
 85 90 95

35 Ser Leu Phe Ala Pro Val Cys Leu Asp Asp Leu Asp Glu Thr Ile Gln
 100 105 110

40 Pro Cys His Ser Leu Cys Val Gln Val Lys Asp Arg Cys Ala Pro Val
 115 120 125

Met Ser Ala Phe Gly Phe Pro Trp Pro Asp Met Leu Glu Cys Asp Arg
 130 135 140

45 Phe Pro Gln Asp Asn Asp Leu Cys Ile Pro Leu Ala Ser Ser Asp His
 145 150 155 160

Leu Leu Pro Ala Thr Glu Glu Ala Pro Lys Val Cys Glu Ala Cys Lys
 165 170 175

50 Asn Lys Asn Asp Asp Asp Asn Asp Ile Met Glu Thr Leu Cys Lys Asn
 180 185 190

55 Asp Phe Ala Leu Lys Ile Lys Val Lys Glu Ile Thr Tyr Ile Asn Arg
 195 200 205

Asp Thr Lys Ile Ile Leu Glu Thr Lys Ser Lys Thr Ile Tyr Lys Leu
 210 215 220

60 Asn Gly Val Ser Glu Arg Asp Leu Lys Lys Ser Val Leu Trp Leu Lys

342

228 230 232 240
 Asp Ser Leu Gln Cys Thr Cys Glu Glu Met Asn Asp Ile Asn Ala Pro
 242 250 258
 5 Tyr Leu Val Met Gly Gln Lys Gln Gly Gly Glu Leu Val Ile Thr Ser
 260 268 276
 10 Val Lys Arg Trp Gln Lys Gly Gln Arg Glu Phe Lys Arg Ile Ser Arg
 278 286 294
 Ser Ile Arg Lys Leu Gln Cys Xaa
 296 298

15

(2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:
 20 (A) LENGTH: 92 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238

25 Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr
 1 5 10 15
 Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp
 20 25 30
 30 Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe
 35 40 45
 35 Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Cys Lys Phe
 50 55 60
 Asn Gln Gln Gln Arg Thr Gly Glu Pro Asp Glu Glu Glu Gly Thr Phe
 65 70 75 80
 40 Arg Ser Ser Ile Arg Arg Leu Ser Xaa Arg Xaa Arg
 85 90

45 (2) INFORMATION FOR SEQ ID NO: 239:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 71 amino acids
 50 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Phe Leu Phe Ile
 1 5 10 15
 55 Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro
 20 25 30
 60 Phe Pro Pro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Ser
 35 40 45

345

Pro Pro Leu Pro Pro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro
 50 55 60

5 Ser Arg Gly Cys Val Leu Leu
 65 70

10 (2) INFORMATION FOR SEQ ID NO: 240:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 71 amino acids

(E) TYPE: amino acid

15 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

Met Pro Gly Thr Phe Leu Arg Pro Phe Val Phe Leu Phe Leu Phe Ile
 1 10 15

20 Cys Cys Cys Leu His Ser Gly Gly Leu Gly Gly Val Pro Leu Pro Pro
 20 25 30

25 Phe Pro Pro Gln Ala Gln Arg Gly Glu Gly Pro Gly Lys Trp Met Ser
 35 40 45

Pro Pro Leu Pro Pro His Pro Val Val Ala Pro Pro Thr Pro Ser Pro
 50 55 60

30 Ser Arg Gly Cys Val Leu Leu
 65 70

35 (2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 amino acids

(E) TYPE: amino acid

40 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

Met Phe Tyr Val Leu Ser Val Ser Xaa Leu Xaa Leu Phe Leu Ala Cys
 1 10 15

45 Gly Leu Cys Leu Xaa Leu Leu Thr Gly Lys Leu Leu
 20 25

50

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 amino acids

55 (E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

60 Met Lys Leu Phe Asp Ala Ser Pro Thr Phe Phe Ala Phe Leu Leu Gly
 1 10 15

34-

His Ile Leu Ala Met Glu Val Leu Ala Trp Leu Leu Ile Tyr Leu Leu
 20 25 30
 5 Gly Pro Gly Trp Val Pro Ser Ala Leu Xaa Arg Leu His Pro Gly His
 35 40 45
 Leu Ser Gly Ser Val Leu Val Ser Ala Ala
 50 55

10

(2) INFORMATION FOR SEQ ID NO: 243:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

20

Met Ile Leu Gly Gly Ile Val Val Val Leu Val Phe Thr Gly Phe Val
 1 5 10 15

25

Trp Ala Ala His Asn Lys Asp Val Leu Arg Arg Met Lys Lys Arg Tyr
 20 25 30

Pro Thr Thr Phe Val Met Val Val Met Leu Ala Ser Tyr Phe Leu Ile
 35 40 45

30

Ser Met Phe Gly Gly Val Met Val Phe Val Phe Gly Ile Thr Phe Pro
 50 55 60

Leu Leu Leu Met Phe Ile His Ala Ser Leu Arg Leu Arg Asn Leu Lys
 65 70 75 80

35

Asn Lys Leu Glu Asn Lys Met Glu Gly Ile Gly Leu Lys Arg Thr Pro
 85 90 95

Met Gly Ile Val Leu Asp Ala Leu Glu Gln Gln Glu Glu Gly Ile Asn
 100 105 110

40

Arg Leu Thr Asp Tyr Ile Ser Lys Val Lys Glu
 115 120

45

(2) INFORMATION FOR SEQ ID NO: 244:

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

55

Ala Leu Val Ser Gly Gln Leu Cys Met Glu Ile Ala Arg Gly Asn Ile
 1 5 10 15

Phe Phe Leu Asn Xaa Leu Val Thr Thr Phe Cys Cys Ser Cys Leu Leu
 20 25 30

60

345

Leu Ser Val Xaa Tyr Leu His Xaa Gly Phe Phe Tyr Ser Ser Leu Cys
 35 40 45

5 Lys Cys Cys Phe Val Leu Val Val Leu Ser Arg Ile Gly Ser Val Asn
 50 55 60

Glu Thr Trp Ser Cys Asn Phe Ser Ile
 65 70

10

(2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:
 15 (A) LENGTH: 49 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

20 Thr Pro Ala Thr Thr Ser Ser Ser Ser Ser Pro Leu Phe Leu Ser Ser
 1 5 10 15

Pro Asp Trp Ser Ser Cys Pro Ser Gly Ser Cys Ile Ala Pro Trp Cys
 20 25 30

25 Thr His Trp Ser Ser Ile Leu Pro Ser Leu Xaa Ile Thr Ser Ser Ile
 35 40 45

30 Pro

35 (2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

Met Ala Arg Val Pro Pro Leu Ser Ser Ser Trp Thr Ser Ser Arg Tyr
 1 5 10 15

45 Arg Arg Trp Leu Cys Cys Pro Val Trp Trp Thr Thr Phe Trp Ala Thr
 20 25 30

Ala Trp Ser Leu Thr Lys His Leu Tyr Lys Asp Val Thr Asp Ala Ile
 35 40 45

50 Arg Asp Val His Val Lys Gly Leu Met Tyr Gln Trp Ile Glu Gln Asp
 50 55 60

55 Met Glu Lys Tyr Ile Leu Arg Gly Asp Glu Thr Phe Ala Val Leu Ser
 65 70 75 80

Arg Leu Val Ala His Gly Lys Gln Leu Phe Leu Ile Thr Asn Ser Pro
 85 90 95

60 Phe Ser Phe Val Asp Lys Gly Met Arg His Met Val Gly Pro Asp Trp

346

100 100 110
 Arg His Ser Ser Met Trp Ser Leu Ser Arg Gln Thr Ser Pro Ala Ser
 110 120 125
 5 Ser Leu Thr Gly Ala Thr Phe Arg Lys Leu Asp Glu Lys Gly Ser Leu
 130 135 140
 Gln Trp Asp Arg Ile Thr Arg Leu Glu Lys Gly Lys Ile Tyr Arg Gln
 140 150 155 160
 Gly Asn Leu Phe Asp Phe Leu Arg Leu Thr Glu Trp Arg Gly Pro Arg
 160 170 175
 15 Val Leu Tyr Phe Gly Asp His Leu Tyr Ser Asp Leu Ala Asp Leu Met
 180 185 190
 Leu Arg His Gly Trp Arg Thr Gly Ala Ile Ile Pro Glu Leu Glu Arg
 190 200 205
 20 Glu Ile Arg Ile Ile Asn Thr Glu Gln Tyr Met His Ser Leu Thr Trp
 210 220 225
 Gln Gln Ala Leu Thr Gly Leu Leu Glu Arg Met Gln Thr Tyr Gln Asp
 225 230 235 240
 Ala Glu Ser Arg Gln Val Leu Ala Ala Trp Met Lys Glu Arg Gln Glu
 240 250 255
 25 Leu Arg Cys Ile Thr Lys Ala Leu Phe Asn Ala Gln Phe Gly Ser Ile
 260 265 270
 Phe Arg Thr Phe His Asn Pro Thr Tyr Phe Ser Arg Arg Leu Val Arg
 270 280 285
 30 Phe Ser Asp Leu Tyr Met Ala Ser Leu Ser Cys Leu Leu Asn Tyr Arg
 290 295 300
 Val Asp Phe Thr Phe Tyr Pro Arg Arg Thr Pro Leu Gln His Glu Ala
 300 310 315 320
 Pro Leu Trp Met Asp Gln Leu Leu His Arg Leu His Glu Asp Pro Leu
 320 330 335
 35 Pro Trp Xaa
 40
 45
 50

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

Met Ala Leu Leu Ser Cys Val Val Asp Tyr Phe Leu Gly His Ser Leu
 1 5 10 15

60

34

Xaa Val.

5

(2) INFORMATION FOR SEQ ID NO: 248:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 339 amino acids

10

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

15

Met Asn Trp Glu Leu Leu Leu Trp Leu Leu Val Leu Cys Ala Leu Leu
1 5 10 15Leu Leu Leu Val Gln Leu Leu Arg Phe Leu Arg Ala Asp Gly Asp Leu
20 25 30

20

Thr Leu Leu Trp Ala Glu Trp Gln Gly Arg Arg Pro Glu Trp Glu Leu
35 40 45Thr Asp Met Val Val Trp Val Thr Gly Ala Ser Ser Gly Ile Gly Glu
50 55 60

25

Glu Leu Ala Tyr Gln Leu Ser Lys Leu Gly Val Ser Leu Val Leu Ser
65 70 75 80

30

Ala Arg Arg Val His Glu Leu Glu Arg Val Lys Arg Arg Cys Leu Gln
85 90 95Asn Gly Asn Leu Lys Glu Lys Asp Ile Leu Val Leu Pro Leu Asp Leu
100 105 110

35

Thr Asp Thr Gly Ser His Glu Ala Ala Thr Lys Ala Val Leu Gln Gly
115 120 125Phe Gly Arg Ile Asp Ile Leu Val Asn Asn Gly Gly Met Ser Gln Arg
130 135 140

40

Ser Leu Cys Met Asp Thr Ser Leu Asp Val Tyr Arg Lys Leu Ile Glu
145 150 155 160

45

Leu Asn Tyr Leu Gly Thr Val Ser Leu Thr Lys Cys Val Leu Pro His
165 170 175Met Ile Glu Arg Lys Gln Gly Lys Ile Val Thr Val Asn Ser Ile Leu
180 185 190

50

Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala Ser Lys His
195 200 205Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg Thr Glu Leu Ala Thr Tyr
210 215 220

55

Pro Gly Ile Ile Val Ser Asn Ile Cys Pro Gly Pro Val Gln Ser Asn
225 230 235 240

60

Ile Val Glu Asn Ser Leu Ala Gly Glu Val Thr Lys Thr Ile Gly Asp
245 250 255

348

Asn Gly Asp Gln Ser His Lys Met Thr Thr Ser Arg Cys Val Arg Leu
 260 265 270
 5 Met Leu Ile Ser Met Ala Asn Asp Leu Lys Glu Val Trp Ile Ser Glu
 275 280 285
 Gln Pro Phe Leu Leu Val Thr Tyr Leu Trp Gln Tyr Met Pro Thr Trp
 290 295 300
 10 Ala Trp Trp Ile Thr Asn Lys Met Gly Lys Lys Arg Ile Glu Asn Phe
 305 310 315 320
 Lys Ser Gly Val Asp Ala Asp Ser Ser Tyr Phe Lys Ile Phe Lys Thr
 325 330 335
 15 Lys His Asp

20

(1) INFORMATION FOR SEQ ID NO: 249:

- (i) SEQUENCE CHARACTERISTICS:
 25 (A) LENGTH: 96 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

30 Met Gly Ala Arg Pro Gly Gly His Pro Gln Lys Trp Ser Phe Leu Trp
 5 10 15
 Ser Leu Ala Leu Trp Leu Pro Leu Ala Leu Ser Val Ser Leu Phe Leu
 20 25 30
 35 Gly Leu Ser Leu Ser Pro Pro Gln Pro Gly Leu Ser Leu Trp Cys Thr
 35 40 45
 Leu Ser Tyr Cys Cys Glu Gln Trp Lys Phe Lys Gly Thr Pro Ser Pro
 40 50 55 60
 Ala Leu Leu Asn Leu Gly Thr Gln Pro Lys Lys Asp Lys Lys Leu Glu
 65 70 75 80
 45 Asp Ser Ile Ala Thr Gln Leu Arg Xaa Leu Pro Glu Lys Asn Ser Asn
 85 90 95

50

(2) INFORMATION FOR SEQ ID NO: 250:

- 55 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 79 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

60

349

Met Ala Leu Thr Phe Leu Leu Val Leu Leu Thr Leu Ala Thr Leu Cys
 1 5 10 15

Thr Arg Leu His Arg Asn Phe Arg Arg Gly Glu Ser Ile Tyr Trp Gly
 5 20 25 30

Pro Thr Ala Asp Ser Gln Asp Thr Val Ala Ala Val Leu Lys Arg Arg
 35 40 45

Leu Leu Gln Pro Ser Arg Arg Val Lys Arg Ser Arg Arg Pro Xaa
 10 50 55 60

Xaa Pro Pro Thr Pro Asp Ser Gly Pro Glu Gly Glu Ser Ser Glu
 65 70 75

15

(2) INFORMATION FOR SEQ ID NO: 251:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 354 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

25 Met Gly Pro Ser Thr Pro Leu Leu Ile Leu Phe Leu Leu Ser Trp Ser
 1 5 10 15

Gly Pro Leu Gln Gly Gln Gln His His Leu Val Glu Tyr Met Glu Arg
 30 20 25 30

Arg Leu Ala Ala Leu Glu Glu Arg Leu Ala Gln Cys Gln Asp Gln Ser
 35 35 40 45

Ser Arg His Ala Ala Glu Leu Arg Asp Phe Lys Asn Lys Met Leu Pro
 50 55 60

Leu Leu Glu Val Ala Glu Lys Glu Arg Glu Ala Leu Arg Thr Glu Ala
 65 70 75 80

40 Asp Thr Ile Ser Gly Arg Val Asp Arg Leu Glu Arg Glu Val Asp Tyr
 85 90 95

Leu Glu Thr Gln Asn Pro Ala Leu Pro Cys Val Glu Phe Asp Glu Lys
 45 100 105 110

Val Thr Gly Gly Pro Gly Thr Lys Gly Lys Gly Arg Arg Asn Glu Lys
 115 120 125

50 Tyr Asp Met Val Thr Asp Cys Gly Tyr Thr Ile Ser Gln Val Arg Ser
 130 135 140

Met Lys Ile Leu Lys Arg Phe Gly Gly Pro Ala Gly Leu Trp Thr Lys
 145 150 155 160

55 Asp Pro Leu Gly Gln Thr Glu Lys Ile Tyr Val Leu Asp Gly Thr Glu
 165 170 175

Asn Asp Thr Ala Phe Val Phe Pro Arg Leu Arg Asp Phe Thr Leu Ala
 60 180 185 190

251

[illegible]

34

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 109 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252

45	Met	Leu	Cys	Ile	Asn	Gly	Thr	Thr	Pro	Arg	Pro	Leu	Pro	Val	Pro	Ser
	1				5					10						15
	Pro	Phe	Gly	Cys	Met	Ile	Phe	Phe	Phe	Phe	Lys	Asn	Pro	Trp	Lys	Gln
				20					25					30		
50	Arg	Leu	Leu	Gln	Gly	Trp	Leu	Gly	Ala	Arg	Pro	Ile	His	Leu	Leu	Gly
			35				40						45			
	Tyr	Leu	Pro	Leu	Ser	Leu	Leu	Trp	Cys	Pro	Phe	Pro	Leu	Pro	Cys	Ala
55		50					55					60				
	Arg	Cys	Ser	Val	Val	Tyr	Ile	Ser	Ser	Pro	Arg	His	Gly	Ala	His	Ala
	65					70					75					80
60	Pro	Arg	Asp	Met	Ile	Leu	Ser	Leu	Val	Leu	Ala	His	Gly	Ala	Leu	Tyr

351

81

90

95

Lys Glu Leu Gly Gly Arg Gly Arg Lys Trp Glu Pro Ser
100 105

5

(2) INFORMATION FOR SEQ ID NO: 253:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

15

Met Phe Tyr Phe Leu Pro Leu Ile Phe Pro Ala Phe Pro Pro Trp Ala
1 5 10 15

20

Phe Arg Leu Ser Thr Leu Phe Thr Ile Ile Ser Trp Ser Glu Asp Ser
20 25 30

Asn Asn Ser Gln Val Tyr Met Asn Cys Val Cys Ser Phe
35 40 45

25

(2) INFORMATION FOR SEQ ID NO: 254:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 315 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

35

Met Ala Gly Gly Arg Cys Gly Pro Xaa Leu Thr Ala Leu Leu Ala Ala
1 5 10 15

Trp Ile Ala Ala Val Ala Ala Thr Ala Gly Pro Glu Glu Ala Ala Leu
20 25 30

40

Pro Pro Glu Gln Ser Arg Val Gln Pro Met Thr Ala Ser Asn Trp Thr
35 40 45

45

Leu Val Met Glu Gly Glu Trp Met Leu Lys Phe Tyr Ala Pro Trp Cys
50 55 60

Pro Ser Cys Gln Gln Thr Asp Ser Glu Trp Glu Ala Phe Ala Lys Asn
65 70 75 80

50

Gly Glu Ile Leu Gln Ile Ser Val Gly Lys Val Asp Val Ile Gln Glu
85 90 95

Pro Gly Leu Ser Gly Arg Phe Phe Val Thr Thr Leu Pro Ala Phe Phe
100 105 110

55

His Ala Lys Asp Gly Ile Phe Arg Arg Tyr Arg Gly Pro Gly Ile Phe
115 120 125

60

Glu Asp Leu Gln Asn Tyr Ile Leu Glu Lys Lys Trp Gln Ser Val Glu
130 135 140

351

Pro Leu Thr Gly Trp Lys Ser Pro Ala Ser Leu Thr Met Ser Gly Met
 14: 15: 16:

5 Ala Gly Leu Phe Ser Ile Ser Gly Lys Ile Trp His Leu His Asn Tyr
 16: 17: 18:

Phe Thr Val Thr Leu Gly Ile Pro Ala Trp Cys Ser Tyr Val Phe Phe
 19: 20:

10 Val Ile Ala Thr Leu Val Phe Gly Leu Phe Met Gly Leu Val Leu Val
 21: 22: 23:

Val Ile Ser Glu Cys Phe Tyr Val Pro Leu Pro Arg His Leu Ser Glu
 24: 25: 26:

Arg Ser Glu Gln Asn Arg Arg Ser Glu Glu Ala His Arg Ala Glu Gln
 27: 28: 29: 30:

20 Leu Gln Asp Ala Glu Glu Lys Asp Asp Ser Asn Glu Glu Glu Asp
 31: 32: 33:

Lys Asp Ser Leu Val Asp Asp Glu Glu Glu Lys Glu Asp Leu Gly Asp
 34: 35: 36:

25 Glu Asp Glu Ala Glu Glu Glu Glu Glu Glu Asp Asn Leu Ala Ala Gly
 37: 38: 39:

Val Asp Glu Glu Arg Ser Glu Ala Asn Asp Gln Gly Pro Pro Gly Gln
 40: 41: 42:

30 Asp Gly Val Thr Arg Glu Xaa Ser Arg Ala Xaa
 43: 44: 45:

35

(2) INFORMATION FOR SEQ ID NO: 255:

- 40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 53 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

45 Met Leu Lys Ala Leu Phe Arg Thr Leu Gln Ala Met Leu Leu Gly Val
 1 5 10 15

Trp Ile Leu Leu Leu Ala Ser Leu Ala Pro Leu Trp Leu Tyr Cys
 20 25 30

50 Trp Arg Met Phe Pro Thr Lys Gly Lys Arg Asp Gln Lys Glu Met Leu
 35 40 45

55 Glu Val Ser Gly Ile
 50

60 (2) INFORMATION FOR SEQ ID NO: 256:

355

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 93 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

Met Ile His Leu Gly His Ile Leu Phe Leu Leu Leu Leu Pro Val Ala
1 10 15

10 Ala Ala Gln Thr Thr Pro Gly Glu Arg Ser Ser Leu Pro Ala Phe Tyr
20 25 30

Pro Gly Thr Ser Gly Ser Cys Ser Gly Cys Gly Ser Leu Ser Leu Pro
35 40 45

15 Leu Leu Ala Gly Leu Val Ala Ala Asp Ala Val Ala Ser Leu Leu Ile
50 55 60

20 Val Gly Ala Val Phe Leu Cys Ala Arg Pro Arg Arg Ser Pro Ala Gln
65 70 75 80

Asp Gly Lys Val Tyr Ile Asn Met Pro Gly Arg Gly Xaa
85 90

25

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

35 Pro Gly His Leu Leu Pro His Lys Trp Glu Asn Cys
1 5 10

40 (2) INFORMATION FOR SEQ ID NO: 258:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1852 base pairs

(B) TYPE: nucleic acid

45 (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

50 TGGCATCTGT GAGCAGTTCG CAGGCTCCGG CCAGGATCCC TTCTTCTCC TCATGGCTC 60

ATGGATCCCA AGGGGCTCCT CTCCTTGACC TTCTGTCTGT TTCTCTCCCT GGCTTTTGG 120

GCAAGCTACG GAACAGGTGG GCGCATGATG AACTGCCCAA AGATTCTCCG GCAATTGGGA 180

55 AGCAAGTGC TGTGCCCCCT GACATATGAA AGGATAAATA AGAGCATGAA CAAAAGCAT 240

CACATTGTCT TCACAATGGC AAAATCACTG GAGAACAGTG TCGAGAACAA AATAGTGTCT 300

60 CTTGATCCAT CGAAGCAGG CCGTCCACGT TATCTAGGAG ATCGCTACAA GTTTTATCT 360

GAGGAGTCTT CTTGGGGGAC ATGAGGAGG ATGAGGAGG GAGGAGTCTT 420
 5 ACGCTGGAGA AAAATGTTTC ATTGAGGCTT TTTGCTGC AGTGGAGGCT TTATGAGGAG 480
 GTTTCAGCTT CAGAAATTAA ATTCTTAAAT AAGAGTCAGG AGAATGGGAC CTACAGCTTT 540
 ATATGAGCTT GACAGTGA GAAGGAGAA GAGTGGCTT ACGCTGGAG TGAAGAAGG 600
 10 GGAAGCAGCT CACTGAACCT ATTCAGAGG TCCAGCTCC TGTGCTCAG CTTGGAGCT 660
 CAGGAGTCTT ACGAGTCTT CAGCTGAGG GTGAGGAGG CTATCAGTAA CAGTGGCAG 720
 ACGTTCAGCT GTGGGCTGG ATGAGGAGG GAGGCTCAG AAACAAAAC ATGGGAGT 780
 15 TATGCTGAGT TGTAGGGGAG TGTGAGTCTT ATTGAGTCTT TGTGAGTCTT ACTAGCTT 840
 AGAGAGAGT STRAAGGAG CAGTCTGAG AAGAGCTGG AAAAAAAG CTTAGCTT 900
 20 TATGCTGAG TGTAGAGAGT ATGAGAGT GAGTCTGAG CTTGGAGCTT ATTAGCTT 960
 AGAGAGAGT TATGCTGAG TGTGAGTCTT GAGTCTGAG AAGAGCTCTT TGTAGAGT 1020
 GTAGAGTCTT GAGTCTGAG GAGGAGTCTT AAGAGCTCTT GAGGAGTCTT 1080
 25 CAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1140
 GAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1200
 30 ATAGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1260
 GAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1320
 CAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1380
 35 ATAGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1440
 AAGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1500
 40 GAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1560
 GAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1620
 TATGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1680
 45 TATGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1740
 ATAGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1800
 50 GAGGAGTCTT CAGGAGTCTT GAGGAGTCTT GAGGAGTCTT GAGGAGTCTT 1860

55 (2) INFORMATION FOR SEQ ID NO: 289:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 371 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

60

355

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

Met Glu Leu Glu Leu Asp Ala Gly Asp Gln Asp Leu Leu Ala Phe Leu
 1 5 10 15
 5 Leu Glu Glu Ser Gly Asp Leu Gly Thr Ala Pro Asp Glu Ala Val Arg
 20 25 30
 10 Ala Pro Leu Asp Trp Ala Leu Pro Leu Ser Glu Val Pro Ser Asp Trp
 35 40 45
 Glu Val Asp Asp Leu Leu Cys Ser Leu Leu Ser Pro Pro Ala Ser Leu
 50 55 60
 15 Asn Ile Leu Ser Ser Ser Asn Pro Cys Leu Val His His Asp His Thr
 65 70 75 80
 Tyr Ser Leu Pro Arg Glu Thr Val Ser Met Asp Leu Glu Ser Glu Ser
 85 90 95
 20 Cys Arg Lys Glu Gly Thr Gln Met Thr Pro Gln His Met Glu Glu Leu
 100 105 110
 25 Ala Glu Gln Glu Ile Ala Arg Leu Val Leu Thr Asp Glu Glu Lys Ser
 115 120 125
 Leu Leu Glu Lys Glu Gly Leu Ile Leu Pro Glu Thr Leu Pro Leu Thr
 130 135 140
 30 Lys Thr Glu Glu Gln Ile Leu Lys Arg Val Arg Arg Lys Ile Arg Asn
 145 150 155 160
 Lys Arg Ser Ala Gln Glu Ser Arg Arg Lys Lys Lys Val Tyr Val Gly
 165 170 175
 35 Gly Leu Glu Ser Arg Val Leu Lys Tyr Thr Ala Gln Asn Met Glu Leu
 180 185 190
 40 Gln Asn Lys Val Gln Leu Leu Glu Glu Gln Asn Leu Ser Leu Leu Asp
 195 200 205
 Gln Leu Arg Lys Leu Gln Ala Met Val Ile Glu Ile Ser Asn Lys Thr
 210 215 220
 45 Ser Ser Ser Ser Thr Cys Ile Leu Val Leu Leu Val Ser Phe Cys Leu
 225 230 235 240
 Leu Leu Val Pro Ala Met Tyr Ser Ser Asp Thr Arg Gly Ser Leu Pro
 245 250 255
 50 Ala Glu His Gly Val Leu Ser Arg Gln Leu Arg Ala Leu Pro Ser Glu
 260 265 270
 55 Asp Pro Tyr Gln Leu Glu Leu Pro Ala Leu Gln Ser Glu Val Pro Lys
 275 280 285
 Asp Ser Thr His Gln Trp Leu Asp Gly Ser Asp Cys Val Leu Gln Ala
 290 295 300
 60 Pro Gly Asn Thr Ser Cys Leu Leu His Tyr Met Pro Gln Ala Pro Ser

356

[illegible]

15

(2) INFORMATION FOR SEQ ID NO: 260:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260

25 Cys Arg Cys Ala Ser Gly Phe Thr Gly Glu Asp Cys
 ; 10

30 (2) INFORMATION FOR SEQ II NC: 261:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11 amino acids
 (E) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261

Cys Thr Cys Gln Val Gly Phe Thr Gly Lys Glu Cys
 1 10

(2) INFORMATION FOR SEC II NO: 262:

4⁵ (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

50 Cys Leu Asn Leu Pro Gly Ser Tyr Gln Cys Gln Cys

55

(2) INFORMATION FOR SEQ ID NO: 263:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid

357

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

Cys Lys Cys Leu Thr Gly Phe Thr Gly Gln Lys Cys
 1 5 10

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

Cys Gln Cys Leu Gln Gly Phe Thr Gly Gln Tyr Cys
 1 5 10

(2) INFORMATION FOR SEQ ID NO: 265:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 127 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

Gly Leu Ala Cys Trp Leu Ala Gly Val Ile Phe Ile Asp Arg Lys Arg
 1 5 10 15

Thr Gly Asp Ala Ile Ser Val Met Ser Glu Val Ala Gln Thr Leu Leu
 20 25 30

Thr Gln Asp Val Xaa Val Trp Val Phe Pro Glu Gly Thr Arg Asn His
 35 40 45

Asn Gly Ser Met Leu Pro Phe Lys Arg Gly Ala Phe His Leu Ala Val
 50 55 60

Gln Ala Gln Val Pro Ile Val Pro Ile Val Met Ser Ser Tyr Gln Asp
 65 70 75 80

Phe Tyr Cys Lys Lys Glu Arg Arg Phe Thr Ser Gly Gln Cys Gln Val
 85 90 95

Arg Val Leu Pro Pro Val Pro Thr Glu Gly Leu Thr Pro Asp Asp Val
 100 105 110

Pro Ala Leu Ala Asp Arg Val Arg His Ser Met Leu His Cys Phe
 115 120 125

(2) INFORMATION FOR SEQ ID NO: 266:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 98 amino acids

(B) TYPE: amino acid

(E) TOPOLOGY: linear

(ix), SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Pro Ser Ala Lys Tyr Phe Phe Lys Met Ala Phe Tyr Asn Gly Trp Ile
10

Leu Phe Leu Ala Val Leu Ala Ile Pro Val Cys Ala Val Arg Gly Arg
20

Asn Val Glu Asn Met Lys Ile Leu Arg Leu Met Leu Leu His Ile Lys
30

Tyr Leu Tyr Gly Ile Arg Val Glu Val Arg Gly Ala His His Phe Pro
40

Pro Ser Gln Pro Tyr Val Val Val Ser Asn His Gln Ser Ser Leu Asp
50

Leu Leu Gly Met Met Glu Val Leu Pro Gly Arg Cys Val Pro Ile Ala
60

Lys Arg

25

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS :

30

(A) LENGTH: 9 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267.

35 Thr Val Phe Arg Glu Ile Ser Thr Asp

40

(2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids

(E) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

Leu Trp Ala Gly Ser Ala Gly Trp Pro Ala Gly
1 10

(2) INFORMATION FOR SEQ ID NO: 269:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

60

359

Ser Ile Leu Gly Ile Ile Ser Val Pro Leu Ser Ile Gly Tyr Cys Ala
 1 5 10 15

5 Ser Lys His Ala Leu Arg Gly Phe Phe Asn Gly Leu Arg
 20 25

10 (2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

Met Ala Tyr His Gly Leu Thr Val
 1 5

20

(2) INFORMATION FOR SEQ ID NO: 271:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

30 Ile Ser Ala Ala Arg Val
 1 5

35 (2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

Pro Asp Val Ser Glu Phe Met Thr Arg Leu Phe
 1 5 10

45

(2) INFORMATION FOR SEQ ID NO: 273:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 17 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

55 Phe Asp Pro Val Arg Val Asp Ile Thr Ser Lys Gly Lys Met Arg Ala
 1 5 10 15

Arg


360

Applicants or agents file reference number	5001PC	International application	Unassigned
--	--------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 64, line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution: American Type Culture Collection	
Address of depositary institution (including postal code and country): 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97901
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	


For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application.	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

361

Applicant's or agent's file reference number:	US001PC	International application	Unassigned
---	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 134a)

A. The indications made below relate to the microorganism referred to in the description on page 64 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97898
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau or
Authorized officer 	Authorized officer

Applicants or agents file reference number:	5001PC	362	International designation:	Unassigned
---	--------	-----	----------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)


A. The indications made below relate to the microorganism referred to in the description. on page 64, line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209044
C. ADDITIONAL INDICATIONS (leave blank if not applicable). This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g. "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

363

Applicant's or agent's file reference number	US001PC	International application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 136a)


A. The indications made below relate to the microorganism referred to in the description: on page <u>64</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97899
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application.	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer:

364

Applicants or agents file reference number	US001PC	international application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depository institution American Type Culture Collection	
Address of depository institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209045
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer:

Applicant's or agent's file reference number	PS001PCT	365	international application	Unassigned
--	----------	-----	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)


A. The indications made below relate to the microorganism referred to in the description. on page <u>64</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>9790</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

366

Applicants or agents' reference number	5001PC	International application	Classified
--	--------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 6, line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet	
Name of depository institution American Type Culture Collection	
Address of depository institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 20904
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer: 	Authorized officer:

367	
Applicants or agents file reference number: S001PC	International application: Unassigned

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 65 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution: American Type Culture Collection	
Address of depositary institution (including postal code and country): 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit: April 28, 1997	Accession Number: 209010
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application.	<input type="checkbox"/> This sheet was received by the International Bureau or.
Authorized officer: 	Authorized officer:

Applicant or agent's file reference number	368	International application	Unassigned
--	-----	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 126a)


A. The indications made below relate to the microorganism referred to in the description on page 65 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 29, 1997	Accession Number 209085
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau or
Authorized officer 	Authorized officer

369

Applicant's or agent's file reference number:	5001PC	International application:	Unassigned
---	--------	----------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*ter*)


A. The indications made below relate to the microorganism referred to in the description. on page <u>65</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country): 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97897
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer: 	Authorized officer:

370

Applicant's or agent's file reference number	5001PC	International application	Unassigned
--	--------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 6, line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depository institution American Type Culture Collection	
Address of depository institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209043
C. ADDITIONAL INDICATIONS <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	
E. SEPARATE FURNISHING OF INDICATIONS <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application.	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

Applicants or agents' reference number	5001PCT	371	international application	Unassigned
--	---------	-----	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>1</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit September 4, 1997	Accession Number 209236
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit").	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized office: 	Authorized office:

Applicant's or agent's reference number	8001PCT	377	International application	Classified
---	---------	-----	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)


A. The indications made below relate to the microorganism referred to in the description on page <u>1</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depository institution American Type Culture Collection	
Address of depository institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 29, 1997	Accession Number 209084
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For international bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

373

Applicant's or agent's file reference number	S001PC7	International application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 136is.)


A. The indications made below relate to the microorganism referred to in the description. on page <u>70</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209048</u>
C. ADDITIONAL INDICATIONS <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	
E. SEPARATE FURNISHING OF INDICATIONS <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

374

Applicants or agents file reference number	SOCIETY	International application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM


(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description: on page <u>70</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution: <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country): <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit: <u>February 26, 1997</u>	Accession Number: <u>97902</u>
C. ADDITIONAL INDICATIONS <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	
E. SEPARATE FURNISHING OF INDICATIONS <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau or
Authorized officer: 	Authorized officer:

Applicant's or agent's file reference number:	S001PCT	375	International application:	Unassigned
---	---------	-----	----------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*ter*)


A. The indications made below relate to the microorganism referred to in the description on page 77, line N/A.	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection.	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97903
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application.	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

376

Applicants or agents for reference number	8061PCT	international application	UNASSIGNED
---	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)


A. The indications made below relate to the microorganism referred to in the description on page 55 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209046
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer

377

Applicant's or agent's file reference number	5001PCT	International application	Unassigned	7-12-98
--	---------	---------------------------	------------	---------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description. on page 80, line N/A	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97904
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer:


378

Applicants or agents' reference number	5061PCT	International application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>86</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit May 15, 1997	Accession Number 209051
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer 	Authorized officer


375

Applicant's or agent's file reference number	5001PCT	International application number	Unassigned
--	---------	----------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 136is.)

A. The indications made below relate to the microorganism referred to in the description on page <u>6</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution: <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country): <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20855</u> <u>United States of America</u>	
Date of deposit: <u>April 4, 1997</u>	Accession Number: <u>97976</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit").	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on
Authorized officer: 	Authorized officer:


380

Applicant's or agent's file reference number:	5061PC	International application:	Unassigned
---	--------	----------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description: on page <u>64</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution: <p style="text-align: center;">American Type Culture Collection</p>	
Address of depositary institution (including postal code and country): <p>12301 Parklawn Drive Rockville, Maryland 20852 United States of America</p>	
Date of deposit: <u>May 15, 1997</u>	Accession Number: <u>209047</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer: </p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on</p> <p>Authorized officer:</p>
---	--

What Is Claimed Is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:
 - (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
 - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X, having biological activity;
 - (f) a polynucleotide which is a variant of SEQ ID NO:X;
 - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
 - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
 - (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.
2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.
3. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

8. A method of making a recombinant host cell comprising the isolated nucleic acid molecule of claim 1.

9. A recombinant host cell produced by the method of claim 8.

10. The recombinant host cell of claim 9 comprising vector sequences.

11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;

(c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;

(f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z.

(g) a variant of SEQ ID NO:Y;

(h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

5 12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

10 13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of claim 11.

15 15. A method of making an isolated polypeptide comprising:
(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and
(b) recovering said polypeptide.

20 16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

25 18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

30 (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

35 (a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

- 5 (a) contacting the polypeptide of claim 11 with a binding partner; and
(b) determining whether the binding partner effects an activity of the polypeptide.

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

10 22. A method of identifying an activity in a biological assay, wherein the method comprises:

- (a) expressing SEQ ID NO:X in a cell;
(b) isolating the supernatant;
(c) detecting an activity in a biological assay; and
15 (d) identifying the protein in the supernatant having the activity.

23. The product produced by the method of claim 22.

Applicant's or agent's file reference number	PS001PCT	International application	Unassigned
--	----------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution <i>(including postal code and country)</i> <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97900</u>
C. ADDITIONAL INDICATIONS <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	
E. SEPARATE FURNISHING OF INDICATIONS <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later <i>(specify the general nature of the indications, e.g., "Accession Number of Deposit")</i>	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	---

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PC	International application	Unassigned
--	---------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>65</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209043</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No. Unassigned
--	----------	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209044</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") <div style="height: 150px;"></div>	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application.	
Authorized officer	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer <div style="height: 40px;"></div>	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicants or agent's file reference number	P5001PC7	International application	o. Unassigned
---	----------	---------------------------	---------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>65</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209045</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>in respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No.	Unassigned
--	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209046</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") <div style="height: 100px;"></div>	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	---

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No. Unassigned
--	----------	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209047</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") <div style="height: 100px;"></div>	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

Page 2

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number: PS001PCT	International application No. Unassigned
--	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>76</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209048</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
Empty space for designated states	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	
Empty space for separate furnishing of indications	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by an applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicants or agent's file reference number	PS00/PCT	International application	Unassigned
---	----------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>77</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209049</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer </p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application	Unassigned
--	----------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>80</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 15, 1997</u>	Accession Number <u>209050</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>in respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	---

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No.	Unassigned
--	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description: on page <u>65</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>April 28, 1997</u>	Accession Number <u>209010</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on	
Authorized officer	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by an applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No. Unassigned
--	----------	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 64, line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97901
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS601PC	International application to	Unassigned
--	---------	------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>77</u> . line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97903
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; height: 40px;"> Authorized officer <div style="position: absolute; left: 100px; top: 0px;"> </div> </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; height: 40px;"> Authorized officer </div>
--	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number: PS001PCT	International application no. Unassigned
--	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description. on page 64 line N/A	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution American Type Culture Collection	
Address of depositary institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit February 26, 1997	Accession Number 97898
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	---

NORWAY

AUSTRALIA

FINLAND

Environ Biol Fish (2015) 98:1031–1042

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by an applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No.	Unassigned
--	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>80</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97904</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No.	Unassigned
--	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>72</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>May 29, 1997</u>	Accession Number <u>209084</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") <div style="height: 100px;"></div>	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS


The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No.	Unassigned
--	----------	-------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>64</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97899</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer 	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PCT	International application No	Unassigned
--	----------	------------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>65</u> line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97897</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit") <div style="height: 100px;"></div>	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Authorized officer </div>
---	--

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	75001PCT	International application	Unassigned
--	----------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>82</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>April 4, 1997</u>	Accession Number <u>97976</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
<p>In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).</p>	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

CANADA

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

NORWAY

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

Applicant's or agent's file reference number	PS001PC7	International application	Unassigned
--	----------	---------------------------	------------

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>76</u> , line <u>N/A</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution (including postal code and country) <u>12301 Parklawn Drive</u> <u>Rockville, Maryland 20852</u> <u>United States of America</u>	
Date of deposit <u>February 26, 1997</u>	Accession Number <u>97902</u>
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	
<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	

For International Bureau use only	
<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁶ : C12N 15/12, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, G01N 33/50, 33/53, 33/68, A61K 38/17</p>	A3	<p>(11) International Publication Number: WO 98/39446</p> <p>(43) International Publication Date: 11 September 1998 (11.09.98)</p>																														
<p>(21) International Application Number: PCT/US98/04482</p> <p>(22) International Filing Date: 6 March 1998 (06.03.98)</p> <p>(30) Priority Data:</p> <table style="width: 100%; border-collapse: collapse;"> <tr><td style="width: 30%;">60/040,162</td><td style="width: 40%;">7 March 1997 (07.03.97)</td><td style="width: 30%;">US</td></tr> <tr><td>60/040,333</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/038,621</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,161</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,626</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,334</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,336</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/040,163</td><td>7 March 1997 (07.03.97)</td><td>US</td></tr> <tr><td>60/043,580</td><td>11 April 1997 (11.04.97)</td><td>US</td></tr> <tr><td>60/043,568</td><td>11 April 1997 (11.04.97)</td><td>US</td></tr> </table> <p><i>(Continued on the following page)</i></p> <p>(71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Olney, MD 20832 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US). FISCHER, Carrie, L. [US/US]; 5810 Hall Street, Burke, VA 22015 (US). SOPPET, Daniel, R. [US/US]; 15050 Stillfield Place, Centreville, VA 22020 (US). CARTER, Kenneth, C. [US/US]; 11601 Brandy Hall Lane, North Potomac, MD 20878 (US). BEDNARIK, Daniel, P. [US/US]; 8822 Blue Sea Drive, Columbia, MD 21046 (US). ENDRESS, Gregory, A. [US/US]; 9729 Clagett Farm Drive, Potomac, MD 20854 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). FENG, Ping [CN/US]; 4 Relda Court, Gaithersburg, MD 20878 (US). YOUNG, Paul, E. [US/US]; 122 Beckwith Street, Gaithersburg, MD 20878 (US). GREENE, John, M. [US/US]; 872 Diamond Drive, Gaithersburg, MD 20878 (US). FERRIE, Ann, M. [US/US]; 13203 L Astoria Hill Court, Germantown, MD 20874 (US). DUAN, Roxanne [US/US]; 4541 Fairfield Drive, Bethesda, MD 20814 (US). HU, Jing-Shan [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). GRAVES, Kimberly, A. [US/US]; 12805 Atlantic Avenue, Rockville, MD 20851 (US). OLSEN, Henrik, S. [DK/US]; 182 Kendrick Place #24, Gaithersburg, MD 20878 (US). EBNER, Reinhard [DE/US]; 9906 Shelburne Terrace #316, Gaithersburg, MD 20878 (US). BREWER, Laurie, A. [US/US]; 14920 Mt. Nebo Road, Poolesville, MD 20837 (US). MOORE, Paul, A. [GB/US]; Apartment #104, 1908 Holly Ridge Drive, McLean, VA 22102 (US). SHI, Yanggu [CN/US]; 437 West Side Drive, Gaithersburg, MD 20878 (US). LAFLEUR, David, W. [US/US]; 1615 Q Street, N.W. #807, Washington, DC 20009 (US). LI, Yi [CN/US]; 1247 Lakeside Drive #3034, Sunnyvale, CA 94086 (US). ZENG, Zhizhen [CN/US]; 13950 Saddleview Drive, Gaithersburg, MD 20878 (US). KYAW, Hla [BU/US]; 520 Sugarbush Circle, Frederick, MD 21703 (US).</p> <p>(74) Agents: BROOKES, Anders, A. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 10850 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p>			60/040,162	7 March 1997 (07.03.97)	US	60/040,333	7 March 1997 (07.03.97)	US	60/038,621	7 March 1997 (07.03.97)	US	60/040,161	7 March 1997 (07.03.97)	US	60/040,626	7 March 1997 (07.03.97)	US	60/040,334	7 March 1997 (07.03.97)	US	60/040,336	7 March 1997 (07.03.97)	US	60/040,163	7 March 1997 (07.03.97)	US	60/043,580	11 April 1997 (11.04.97)	US	60/043,568	11 April 1997 (11.04.97)	US
60/040,162	7 March 1997 (07.03.97)	US																														
60/040,333	7 March 1997 (07.03.97)	US																														
60/038,621	7 March 1997 (07.03.97)	US																														
60/040,161	7 March 1997 (07.03.97)	US																														
60/040,626	7 March 1997 (07.03.97)	US																														
60/040,334	7 March 1997 (07.03.97)	US																														
60/040,336	7 March 1997 (07.03.97)	US																														
60/040,163	7 March 1997 (07.03.97)	US																														
60/043,580	11 April 1997 (11.04.97)	US																														
60/043,568	11 April 1997 (11.04.97)	US																														
<p>Published</p> <p><i>With international search report.</i></p> <p><i>With an indication in relation to a deposited microorganism furnished under Rule 13^{bis} separately from the description.</i></p> <p><i>Date of receipt by the International Bureau:</i> 06 April 1998 (06.04.98)</p> <p>(88) Date of publication of the international search report: 23 December 1998 (23.12.98)</p>																																
<p>(54) Title: 70 HUMAN SECRETED PROTEINS</p> <p>(57) Abstract</p> <p>The present invention relates to 70 novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating disorders related to these novel human secreted proteins.</p>																																

60/043,314	11 April 1997 (11.04.97)	US	60/047,598	23 May 1997 (23.05.97)	US	60/056,882	22 August 1997 (22.08.97)	US
60/043,569	11 April 1997 (11.04.97)	US	60/047,613	23 May 1997 (23.05.97)	US	60/056,637	22 August 1997 (22.08.97)	US
60/043,311	11 April 1997 (11.04.97)	US	60/047,582	23 May 1997 (23.05.97)	US	60/056,903	22 August 1997 (22.08.97)	US
60/043,671	11 April 1997 (11.04.97)	US	60/047,596	23 May 1997 (23.05.97)	US	60/056,888	22 August 1997 (22.08.97)	US
60/043,674	11 April 1997 (11.04.97)	US	60/047,612	23 May 1997 (23.05.97)	US	60/056,879	22 August 1997 (22.08.97)	US
60/043,669	11 April 1997 (11.04.97)	US	60/047,632	23 May 1997 (23.05.97)	US	60/056,880	22 August 1997 (22.08.97)	US
60/043,312	11 April 1997 (11.04.97)	US	60/047,601	23 May 1997 (23.05.97)	US	60/056,894	22 August 1997 (22.08.97)	US
60/043,313	11 April 1997 (11.04.97)	US	60/047,595	23 May 1997 (23.05.97)	US	60/056,911	22 August 1997 (22.08.97)	US
60/043,672	11 April 1997 (11.04.97)	US	60/047,599	23 May 1997 (23.05.97)	US	60/056,636	22 August 1997 (22.08.97)	US
60/043,315	11 April 1997 (11.04.97)	US	60/047,588	23 May 1997 (23.05.97)	US	60/056,874	22 August 1997 (22.08.97)	US
60/043,578	11 April 1997 (11.04.97)	US	60/047,585	23 May 1997 (23.05.97)	US	60/056,910	22 August 1997 (22.08.97)	US
60/043,576	11 April 1997 (11.04.97)	US	60/047,586	23 May 1997 (23.05.97)	US	60/056,864	22 August 1997 (22.08.97)	US
60/043,670	11 April 1997 (11.04.97)	US	60/047,590	23 May 1997 (23.05.97)	US	60/056,631	22 August 1997 (22.08.97)	US
60/047,600	23 May 1997 (23.05.97)	US	60/047,594	23 May 1997 (23.05.97)	US	60/056,845	22 August 1997 (22.08.97)	US
60/047,615	23 May 1997 (23.05.97)	US	60/047,589	23 May 1997 (23.05.97)	US	60/056,892	22 August 1997 (22.08.97)	US
60/047,597	23 May 1997 (23.05.97)	US	60/047,593	23 May 1997 (23.05.97)	US	60/056,632	22 August 1997 (22.08.97)	US
60/047,502	23 May 1997 (23.05.97)	US	60/047,614	23 May 1997 (23.05.97)	US	60/056,664	22 August 1997 (22.08.97)	US
60/047,633	23 May 1997 (23.05.97)	US	60/047,501	23 May 1997 (23.05.97)	US	60/056,876	22 August 1997 (22.08.97)	US
60/047,583	23 May 1997 (23.05.97)	US	60/048,974	06 June 1997 (06.06.97)	US	60/056,881	22 August 1997 (22.08.97)	US
60/047,617	23 May 1997 (23.05.97)	US	60/048,964	06 June 1997 (06.06.97)	US	60/056,909	22 August 1997 (22.08.97)	US
60/047,618	23 May 1997 (23.05.97)	US	60/056,886	22 August 1997 (22.08.97)	US	60/056,875	22 August 1997 (22.08.97)	US
60/047,503	23 May 1997 (23.05.97)	US	60/056,877	22 August 1997 (22.08.97)	US	60/056,862	22 August 1997 (22.08.97)	US
60/047,592	23 May 1997 (23.05.97)	US	60/056,889	22 August 1997 (22.08.97)	US	60/056,887	22 August 1997 (22.08.97)	US
60/047,581	23 May 1997 (23.05.97)	US	60/056,893	22 August 1997 (22.08.97)	US	60/056,908	22 August 1997 (22.08.97)	US
60/047,584	23 May 1997 (23.05.97)	US	60/056,630	22 August 1997 (22.08.97)	US	60/056,884	22 August 1997 (22.08.97)	US
60/047,500	23 May 1997 (23.05.97)	US	60/056,878	22 August 1997 (22.08.97)	US	60/057,761	05 September 1997 (05.09.97)	US
60/047,587	23 May 1997 (23.05.97)	US	60/056,662	22 August 1997 (22.08.97)	US	60/057,650	05 September 1997 (05.09.97)	US
60/047,492	23 May 1997 (23.05.97)	US	60/056,872	22 August 1997 (22.08.97)	US			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MX	Mexico	UG	Uganda
BY	Belarus	IS	Iceland	NE	Niger	US	United States of America
CA	Canada	IT	Italy	NL	Netherlands	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NO	Norway	VN	Viet Nam
CG	Congo	KE	Kenya	NZ	New Zealand	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	PL	Poland	ZW	Zimbabwe
CJ	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PT	Portugal		
CM	Cameroon	KR	Republic of Korea	RO	Romania		
CN	China	KZ	Kazakhstan	RU	Russian Federation		
CU	Cuba	LC	Saint Lucia	SD	Sudan		
CZ	Czech Republic	LI	Liechtenstein	SE	Sweden		
DE	Germany	LK	Sri Lanka	SG	Singapore		
DK	Denmark	LR	Liberia				
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No.

PC1/US 98/04482

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N5/10 C12N1/21 C07K14/47 C07K16/18
 C12Q1/68 G01N33/50 G01N33/53 G01N33/68 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>L. HILLIER ET AL.: "The WashU-Merck EST Project 1997"</p> <p>EMBL SEQUENCE DATABASE, 6 March 1997, HEIDELBERG, FRG, XP002068123</p> <p>zr78g10.r1 Soares NhMPu S1 Homo sapiens cDNA clone 669570 5' similar to SW:FUCO_RAT P17164 Alpha-L-fucosidase precursor; Accession. Accession no. AA234924; --- -/--</p>	<p>1-3, 7-10,21</p>

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

Z document member of the same patent family

Date of the actual completion of the international search

16 June 1998

Date of mailing of the international search report

16. 09. 1998

Name and mailing address of the ISA

European Patent Office, P. B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

HORNIG H.

INTERNATIONAL SEARCH REPORT

International Application No

PC1/US 98/04482

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 15 December 1996, HEIDELBERG, FRG, XP002068124 z140b11.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 504381 5' similar to TR:G182779 Lysosomal Enzyme Alpha-L-Fucosidase Accession no. AA151194 ---	1-3, 7-10,21
X	L. HILLIER ET AL.: "The WashU-Merck EST Project" EMBL SEQUENCE DATABASE, 4 June 1996, HEIDELBERG, FRG, XP002068125 zc54a02.r1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 326090 5' similar to SW:FUCO_HUMAN P4066 tissue Alpha-L-Fucosidase precursor; Accession no. W52490 ---	1-3, 7-10,21
A	WO 97 07198 A (GENETICS INSTITUT) 27 February 1997 see the whole document ---	1-23
A	WO 97 04097 A (GENETICS INST) 6 February 1997 see the whole document ---	1-23
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 see the whole document ---	1-23
A	JACOBS K ET AL: "A novel method for isolating eukaryotic cDNA clones encoding secreted proteins." KEYSTONE SYMPOSIUM ON DENDRITIC CELLS: ANTIGEN PRESENTING CELLS OF T AND B LYMPHOCYTES, TAOS, NEW MEXICO, USA, MARCH 10-16, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (21A). 1995. 19. ISSN: 0733-1959, XP002027246 abstract no. C1-207 see abstract ---	1-23
A	WO 90 14432 A (GENETICS INST) 29 November 1990 see the whole document ---	1-23
A	WO 96 17925 A (IMMUNEX CORP) 13 June 1996 see the whole document ---	1-23

-/--

INTERNATIONAL SEARCH REPORT

International Application No.

PC1/US 98/04482

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>T. OCCHIODORO ET AL.: "Human alpha-L-Fucosidase: Complete coding sequence from cDNA clones" BIOCHEM. AND BIOPHYS. RES. COMMUNICATIONS, vol. 164, no. 1, 16 October 1989, ACADEMIC PRESS, NEW YORK, US, pages 439-445, XP002068126 cited in the application see the whole document -----</p>	1-23

INTERNATIONAL SEARCH REPORT

International application No

PCT/US 98/04482

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos. because they relate to subject matter not required to be searched by this Authority, namely
Remark: Although claim 17 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos. because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically
3. ☐ Claims Nos. because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

see further information sheet

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1. Claims: (1-23) partially

-An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence consisting of SEQ ID no. 11; wherein said polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein encoding the sequence of SEQ ID no. 134 or the polypeptide encoded by the cDNA sequence included in ATCC Deposit no: HGCMD20, which is hybridizable to SEQ ID no.11; a recombinant vector comprising said isolated nucleic acid molecule; a method of making a recombinant host cell comprising said isolated nucleic acid molecule; a recombinant host cell comprising said vector; an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence consisting of SEQ ID no. 134; an isolated antibody that binds specifically to said isolated polypeptide; a recombinant host cell that expresses said isolated polypeptide; a method of making said polypeptide; a method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of said polypeptide; a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject using said polynucleotide and/or polypeptide sequences; a method for identifying a binding partner to said polypeptide; a gene corresponding to the cDNA sequences of SEQ ID no.11; a method for identifying an activity in a biological assay, by using the expression of SEQ ID no. 134;

Inventions 2 to 70. Claims: (1-23) partially

-Idem as subject 1 but limited to gene nos. 2 to 70 respectively cDNA clone sequences HLDBG33 to HMCAB89.
(Invention 2 is limited to SEQ ID nos.12,81,135, and 204;
Invention 3 is limited to SEQ ID nos.13 and 136;;
Invention 70 is limited to SEQ ID nos.80 and 203;)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 98/04482

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9707198 A	27-02-97	US 5707829 A	13-01-98
		AU 6712396 A	18-02-97
		AU 6768596 A	12-03-97
		EP 0839196 A	06-05-98
		EP 0851875 A	08-07-98
		WO 9704097 A	06-02-97

WO 9704097 A	06-02-97	US 5707829 A	13-01-98
		AU 6712396 A	18-02-97
		EP 0839196 A	06-05-98
		AU 6768596 A	12-03-97
		EP 0851875 A	08-07-98
		WO 9707198 A	27-02-97

US 5536637 A	16-07-96	US 5712116 A	27-01-98

WO 9014432 A	29-11-90	US 5580753 A	03-12-96
		AT 147436 T	15-01-97
		AU 637620 B	03-06-93
		AU 5928990 A	18-12-90
		CA 2056997 A	24-11-90
		DE 69029657 D	20-02-97
		DK 473724 T	14-04-97
		EP 0473724 A	11-03-92
		ES 2099096 T	16-05-97
		JP 4506006 T	22-10-92
		US 5734037 A	31-03-98
		US 5414071 A	09-05-95

WO 9617925 A	13-06-96	AU 4639396 A	26-06-96
		CA 2206488 A	13-06-96
		FI 972390 A	05-06-97
		NO 972455 A	06-08-97
